



Luis Filipe Mesquita
Nero Moreira Alves

Estudo e Implementação de Amplificadores em
Modo de Corrente com Grande Produto Ganho
Largura de Banda



**Luis Filipe Mesquita
Nero Moreira Alves**

**Estudo e Implementação de Amplificadores em
Modo de Corrente com Grande Produto Ganho
Largura de Banda**

**High Gain and Bandwidth Current-Mode Amplifiers:
Study and Implementation**

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Electrotécnica, realizada sob a orientação científica do Doutor Rui Luís Andrade Aguiar, Professor Auxiliar do Departamento de Electrónica Telecomunicações e Informática da Universidade de Aveiro

o júri

presidente

Doutor Sushil Kumar Mendiratta

Professor Catedrático do Departamento de Física da Universidade de Aveiro

Doutor Dinis Gomes de Magalhães dos Santos

Professor Catedrático do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.

Doutor José Higinio Gomes Correia

Professor Associado com Agregação do Departamento de Electrónica Industrial da Universidade do Minho.

Doutora Guiomar Gaspar de Andrade Evans

Professora Auxiliar do Departamento de Física da Faculdade de Ciências da Universidade de Lisboa

Doutor Rui Luís Andrade Aguiar

Professor Auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.

agradecimentos

Nenhuma secção de agradecimentos é suficientemente extensa de forma a conter as menções a todos aqueles que directa ou indirectamente contribuíram no desenvolvimento e conclusão deste trabalho. Sem pretender minorar o contributo de ninguém, aqui ficam as minhas menções e agradecimentos aos que directamente contribuíram neste trabalho:

Em primeiro lugar ao meu orientador, o Professor Doutor Rui Luís Andrade Aguiar por ter propiciado as condições necessárias à realização deste trabalho, por todo o apoio prestado e por toda a critica científica.

Ao Professor Doutor José Carlos Pedro por muitas e frutuosas conversas que fomos mantendo durante os intervalos de actividade docente.

Aos meus colegas do Laboratório de Circuitos e Sistemas Integrados, em particular aos Mestres Miguel Madureira, Mónica Figueiredo, Eliseu Macedo e Miguel Bergano e ao Licenciado Luis Barbosa, por todo o apoio prestado e pelo ambiente pleno de camaradagem.

Aos meus pais, António e Maria de Fátima por toda a paciência que demonstraram nos períodos mais críticos do desenvolvimento deste trabalho.

Aos meus amigos António, Marta, Magalhães, Célia, Sérgio, Carla, Fátima, Renata, Gustavo, Nuno Gil, Carlos, Luís Mendes, Ivone e tantos outros, pela sua amizade incondicional.

Aos técnicos do Instituto de Telecomunicações Miguel Lacerda, Paulo Gonçalves e Pedro Silva por todo o apoio prestado.

Às instituições de acolhimento onde este trabalho foi desenvolvido, a Universidade de Aveiro e o Instituto de Telecomunicações, por todos os recursos disponibilizados.

Finalmente, mas não menos importante a todos os músicos que tornaram as longas horas de vigília, especialmente durante a fase de escrita, mais suportáveis, em particular, Henry Cow, Van der Graaf Generator, Von Samla (e todas as suas metamorfoses), Can, Aksak Maboul, Pascal Comelade, Laurie Anderson, André Duchesne, Tom Waits, Marianne Faithfull, Serge Gainsbourg, John Coltrane, Ornnette Coleman, Ken Vandermark entre muitos outros.

palavras-chave

Modo de corrente, Amplificadores com retroacção, Produto ganho largura de banda, Atrasos, Malhas de retroacção com atraso, Adaptação activa em modo de corrente.

resumo

Esta tese aborda o problema do projecto de amplificadores com grandes produtos de ganho por largura de banda. A aplicação final considerada consistiu no projecto de amplificadores adequados à recepção de sinais ópticos em sistemas de transmissão ópticos usando o espaço livre. Neste tipo de sistemas as maiores limitações de ganho e largura de banda surgem nos circuitos de entrada. O uso de detectores ópticos com grande área fotosensível é uma necessidade comum neste tipo de sistemas. Estes detectores apresentam grandes capacidades intrínsecas, o que em conjunto com a impedância de entrada apresentada pelo amplificador estabelece sérias restrições no produto do ganho pela largura de banda. As técnicas mais tradicionais para combater este problema recorrem ao uso de amplificadores com retroacção baseados em configurações de transimpedância. Estes amplificadores apresentam baixas impedâncias de entrada devido à acção da retroacção. Contudo, os amplificadores de transimpedância também apresentam uma relação directa entre o ganho e a impedância de entrada. Logo, diminuir a impedância de entrada implica diminuir o ganho. Esta tese propõe duas técnicas novas para combater os problemas referidos. A primeira técnica tem por base uma propriedade fundamental dos amplificadores com retroacção. Em geral, todos os circuitos electrónicos têm tempos de atraso associados, os amplificadores com retroacção não são uma excepção a esta regra. Os tempos de atraso são em geral reconhecidos como elementos instabilizadores neste tipos da amplificadores. Contudo, se usados judiciosamente, este tempos de atraso podem ser explorados como uma forma de aumentar a largura de banda em amplificadores com retroacção. Com base nestas ideias, esta tese apresenta o conceito geral de reatracção com atraso, como um método de optimização de largura de banda em amplificadores com retroacção.

O segundo método baseia-se na destruição da dualidade entre ganho e impedância de entrada existente nos amplificadores de transimpedância. O conceito de adaptação activa em modo de corrente é neste sentido uma forma adequada para separar o detector óptico da entrada do amplificador. De acordo com este conceito, emprega-se um elemento de adaptação em modo de corrente para isolar o detector óptico da entrada do amplificador. Desta forma as tradicionais limitações de ganho e largura de banda podem ser tratadas em separado.

Esta tese defende o uso destas técnicas no desenho de amplificadores de transimpedância para sistemas de recepção de sinais ópticos em espaço livre.

keywords

Current-Mode, Feedback amplifiers, Gain-Bandwidth product, Delays, Delayed Feedback, Active current matching

abstract

This thesis addresses the problem of achieving high gain-bandwidth products in amplifiers. The adopted framework consisted on the design of a free-space optical (FSO) front end amplifier able to amplify very small optical signals over large frequency bandwidths. The major gain-bandwidth limitations in FSO front end amplifiers arise due to the input circuitry. Usually, it is necessary to have large area optical detectors in order to maximize signal reception. These detectors have large intrinsic capacitances, which together with the amplifier input impedance poses a severe restriction on the gain-bandwidth product. Traditional techniques to combat this gain-bandwidth limitation resort to feedback amplifiers consisting on transimpedance configurations. These amplifiers have small input impedances due to the feedback action. Nevertheless, transimpedance amplifiers have a direct relation between gain and input impedance. Thus reducing the input impedance usually implies reducing the gain.

This thesis advances two new methods suitable to combat the above mentioned problems. The first method is based on a fundamental property of feedback amplifiers. In general, all electronic circuits have associated time delays, and feedback amplifiers are not an exception to this rule. Time delays in feedback amplifiers have been recognized as destabilizing elements. Nevertheless, when used with appropriate care, these delays can be exploited as bandwidth enhancement elements. Based on these ideas, this thesis presents the general concept of delayed feedback, as a bandwidth optimization method suitable for feedback amplifiers.

The second method is based on the idea of destroying the impedance-gain duality in transimpedance amplifiers. The concept of active current matching is in this sense a suitable method to detach the optical detector from the transimpedance amplifier input. According to this concept, a current matching device (CMD) is used to convey the signal current sensed by the optical detector, to the amplifier's input. Using this concept the traditional gain-bandwidth limitations can be treated in a separate fashion.

This thesis advocates the usage of these techniques for the design of transimpedance amplifiers suited for FSO receiving systems.

dedicatória

Aos meus afilhados João Miguel e Rafael, para que um dia possam perceber o que está aqui contido e muito mais.
Aos meus pais António e Maria de Fátima por todo o apoio e compreensão.

“...Estes exemplos permitiram que um bibliotecário de génio descobrisse a lei fundamental da Biblioteca. Este pensador observou que todos os livros, por muito diferentes que sejam, constam de elementos iguais: o espaço, o ponto, a vírgula, as vinte e duas letras do alfabeto. Também acrescentou um facto que todos os viajantes têm confirmado: Não há, na vasta Biblioteca, dois livros idênticos. Destas premissas incontroversas deduziu que a Biblioteca é total e que as suas estantes registam todas as possíveis combinações dos vinte e tal símbolos ortográficos (número, embora vastíssimo, não infinito) ou seja, tudo o que nos é dado exprimir: em todos os idiomas. Tudo: a história minuciosa do futuro, as autobiografias dos arcanjos, o catálogo fiel da Biblioteca, milhares e milhares de catálogos falsos, a demonstração da falácia desses catálogos, a demonstração da falácia do catálogo verdadeiro, o evangelho gnóstico de Basilides, o comentário desse evangelho, o comentário do comentário desse evangelho, o relato verídico da tua morte, a versão de cada livro em todas as línguas, as interpolações de cada livro em todos os livros,...”

A Biblioteca de Babel, Jorge Luis Borges.

“...Quando se proclamou que a Biblioteca abrangia todos os livros, a primeira impressão foi de extravagante felicidade. Todos os homens se sentiram senhores de um tesouro intacto e secreto. Não havia problema pessoal ou mundial cuja eloquente solução não existisse, nalgum hexágono. O universo estava justificado, o universo bruscamente usurpou as dimensões ilimitadas da esperança. Naquele tempo falou-se muito das Reabilitações: livros de apologia e de profecia, que para sempre reabilitavam os actos de todos os homens do universo e guardavam arcanos prodigiosos para o seu porvir. Milhares de cobiçosos abandonaram o doce hexágono natal e lançaram-se pelas escadas acima, impelidos pelo vão propósito de encontrar a sua Reabilitação. Estes peregrinos brigavam nos corredores estreitos, proferiam obscuras maldições, estrangulavam-se nas escadas divinas, atiravam os livros enganadores para o fundo dos túneis, morriam defenestrados pelos homens de regiões remotas. Outros enlouqueceram... As Reabilitações existem, (eu vi duas que se referem a pessoas do futuro, a pessoas porventura imaginárias) mas os pesquisadores não se lembravam que a possibilidade de um homem achar a sua, ou alguma pérvida variação da sua, se pode computar à volta de zero.”

A Biblioteca de Babel, Jorge Luis Borges.

Table of Contents

1	INTRODUCTION	1
1.1	PROBLEM STATEMENT	2
1.1.1	FREE-SPACE OPTICAL FRONT-ENDS	2
1.2	METHODS TO COMBAT GBW LIMITATIONS	5
1.3	PROPOSED SOLUTIONS	6
1.3.1	DELAYED FEEDBACK CONCEPT	7
1.3.2	CURRENT MATCHING CONCEPT	8
1.4	THESIS FRAMEWORK	9
1.4.1	WIRELESS INFRARED FSO SYSTEMS	9
1.4.2	VISIBLE LIGHT COMMUNICATION SYSTEMS	10
1.5	SOFTWARE SUPPORT	11
1.6	ORIGINAL CONTRIBUTIONS	12
1.7	CHAPTER OVERVIEW	13
2	GAIN-BANDWIDTH OPTIMIZATION TECHNIQUES	15
2.1	PRELIMINARY CONSIDERATIONS	16
2.2	GENERAL GAIN-BANDWIDTH LIMITATIONS	18
2.2.1	TWO-TERMINAL INTERSTAGES	19
2.2.2	FOUR-TERMINAL INTERSTAGES	21
2.3	CHAIN AMPLIFIERS	23
2.4	STAGGER-TUNED AMPLIFIERS	27
2.5	DISTRIBUTED AMPLIFIERS	31
2.6	FREQUENCY PEAKING	34
2.6.1	BUTTERWORTH DESIGN	34
2.6.2	INDUCTIVE PEAKING	36
2.6.3	T-COIL PEAKING	41
2.6.4	CAPACITIVE PEAKING	43
2.7	FEEDBACK AMPLIFIERS	46
2.7.1	GBW PROPERTIES OF FEEDBACK AMPLIFIERS	46
2.7.2	MAXIMALLY FLAT DESIGN	47
2.7.3	VARIABLE FEEDBACK	49

2.8	CIRCUIT LEVEL TECHNIQUES	54
2.8.1	CASCODE AMPLIFIER	55
2.8.2	CHERRY-HOOPER AMPLIFIER	57
2.8.3	F_T DOUBLER	59
2.8.4	CURRENT-MODE STRATEGIES	60
2.9	CONCLUDING REMARKS	62
3	<u>DELAYED FEEDBACK AMPLIFIERS - THEORY</u>	63
3.1	DELAYED FEEDBACK AMPLIFIERS	64
3.1.1	GENERAL DISCUSSION	67
3.1.2	BOUNDED FREQUENCY RESPONSE LEMMA	73
3.1.3	MAXIMUM BANDWIDTH ENHANCEMENT	76
3.2	DESIGN CONSTRAINTS	78
3.2.1	MAXIMALLY FLAT DESIGN	78
3.2.2	STABILITY LIMITS	82
3.3	DELAYED FEEDBACK AMPLIFIERS TYPES	86
3.3.1	1 ST ORDER DELAYED FEEDBACK AMPLIFIERS	86
3.3.2	2 ND ORDER DELAYED FEEDBACK AMPLIFIERS	88
3.3.3	DELAYED FEEDBACK AMPLIFIERS WITH ALL-PASS DELAYS	90
3.4	ROBUSTNESS ISSUES	96
3.4.1	GAIN AND PHASE MARGINS	97
3.4.2	ROBUSTNESS PROPERTIES UNDER MAXIMAL FLAT CONDITIONS	98
3.5	BANDWIDTH OPTIMIZATION	101
3.5.1	BANDWIDTH ENHANCEMENT FUNCTIONS	101
3.5.2	DELAY OPTIMIZATION	105
3.5.3	GAIN OPTIMIZATION	107
3.6	CONCLUDING REMARKS	108
4	<u>DELAYED FEEDBACK AMPLIFIERS – DESIGN CONSIDERATIONS</u>	109
4.1	GENERAL DISCUSSION	110
4.2	ASYMPTOTIC GAIN MODEL	112
4.2.1	PRELIMINARY DEFINITIONS	113
4.2.2	RETURN RATIO AND RETURN DIFFERENCE	114
4.2.3	ASYMPTOTIC GAIN MODEL FORMULATION	115
4.3	DELAYS IN ACTIVE DEVICES	117

4.3.1	DELAYS IN BIPOLAR TRANSISTORS	119
4.3.2	DELAYS IN FIELD EFFECT TRANSISTORS	123
4.4	DELAY DESIGN	126
4.4.1	GENERAL DELAY REPRESENTATION	126
4.4.2	TRANSMISSION LINE DESIGN	128
4.4.3	ACTIVE DELAY DESIGN	139
4.4.4	ACTIVE DELAY VERSUS TRANSMISSION LINES IN INTEGRATED CIRCUITS	146
4.5	TRANSIMPEDANCE AMPLIFIER DESIGN CONSIDERATIONS	148
4.5.1	CLOSED-LOOP GAIN AND RETURN RATIO	149
4.5.2	POLE AND ZERO CONTRIBUTIONS	152
4.5.3	DELAY CONTRIBUTIONS	155
4.6	CONCLUDING REMARKS	157
5	CURRENT-MODE TECHNIQUES	159
<hr/>		
5.1	CURRENT-MODE DESIGN TECHNIQUES	160
5.1.1	CURRENT-MODE DESIGN PRELIMINARIES	162
5.2	ACTIVE CURRENT-MATCHING CONCEPT	166
5.2.1	REFERENCE TRANSIMPEDANCE AMPLIFIER	166
5.2.2	TIA PLUS ACTIVE CURRENT MATCHING	168
5.2.3	GLOBAL FEEDBACK TIA PLUS CMD	172
5.2.4	DELAYED FEEDBACK	173
5.3	CURRENT CONVEYORS	174
5.3.1	CURRENT CONVEYOR TYPES	175
5.3.2	CURRENT CONVEYOR BASED CMD DESIGN	177
5.3.3	COMMON TOPOLOGIES	179
5.4	POSITIVE CLASS AB CC-II DESIGN CONSIDERATIONS	183
5.4.1	CURRENT DYNAMIC RANGE LIMITATIONS	183
5.4.2	VOLTAGE DYNAMIC RANGE LIMITATIONS	186
5.4.3	FREQUENCY RESPONSE OPTIMIZATION	187
5.4.4	CLASS A HALF CIRCUIT NORMALIZED FREQUENCY RESPONSE	189
5.4.5	COMPLETE FREQUENCY RESPONSE	194
5.4.6	DESIGN GUIDELINES FOR LARGE BANDWIDTH TRANSLINEAR CC-IIs	199
5.5	INPUT IMPEDANCE REDUCTION	201
5.5.1	AVAILABLE TECHNIQUES	201
5.5.2	POSITIVE FEEDBACK	203
5.5.3	NEGATIVE FEEDBACK WITH OA	205

5.5.4	NEGATIVE FEEDBACK WITH TWO CC-IIS	207
5.5.5	PERFORMANCE COMPARISON	209
5.5.6	DUAL-LOOP POSITIVE-NEGATIVE FEEDBACK	211
5.6	CONCLUDING REMARKS	213
6	EXPERIMENTAL VALIDATION	215

6.1	INTRODUCTION	216
6.2	DELAYED FEEDBACK - EXPERIMENTAL VALIDATION	216
6.2.1	EXPERIMENTAL SET-UP	217
6.2.2	NUMERICAL MODEL	218
6.2.3	RESULTS AND DISCUSSION	219
6.3	DELAYED FEEDBACK – INTEGRATED CIRCUIT IMPLEMENTATION	222
6.3.1	REFERENCE TRANSIMPEDANCE AMPLIFIER SPECIFICATIONS	222
6.3.2	OVERALL CIRCUIT DESIGN	223
6.3.3	SIMULATED DELAY CHARACTERIZATION	224
6.3.4	DELAYED FEEDBACK VERIFICATION	227
6.3.5	POST LAYOUT SIMULATION	231
6.4	ACTIVE CURRENT MATCHING CONCEPT VALIDATION	233
6.4.1	SIMULATION SET-UP	233
6.4.2	CCII CIRCUIT DESIGN AND OPTIMIZATION	234
6.4.3	CMD DESIGN OPTIMIZATION	241
6.4.4	CMD PLUS TIA DESIGN OPTIMIZATION	244
6.5	CONCLUDING REMARKS	247
7	CONCLUSIONS	249

7.1	FINAL DISCUSSION	250
7.2	CONCLUSIONS	250
7.2.1	ACTIVE CURRENT MATCHING	250
7.2.2	DELAYED FEEDBACK CONCEPT	251
7.2.3	DELAYED FEEDBACK VERSUS ACTIVE CURRENT MATCHING	251
7.3	GUIDELINES FOR FUTURE WORK	252
APPENDIX		255
A 1	GENERIC TRANSISTOR	256
A 2	BUTTERWORTH POLYNOMIALS	258

B 1	G_{\max} ANALYTICAL APPROXIMATION	259
B 2	THEOREM 3.4 (DEMONSTRATION)	259
B 3	THEOREM 3.5 (DEMONSTRATION)	261
B 4	MIXED DELAYS – STABILITY RESTRICTIONS	262
C 1	MEASUREMENT PROCEDURES	263
C 2	BLACKMAN THEOREM	265
	REFERENCES	267

Table of Figures

FIGURE 1.1 - FSO COMMUNICATION SYSTEM (PHYSICAL LEVEL).	2
FIGURE 1.2 – SMALL SIGNAL EQUIVALENT OF THE INPUT CIRCUITRY OF THE FRONT-END AMPLIFIER.	3
FIGURE 1.3 – TRANSIMPEDANCE FRONT-END.	4
FIGURE 1.4 – DELAYED FEEDBACK AMPLIFIER (CONCEPTUAL DIAGRAM).	7
FIGURE 1.5 – FSO FRONT-END AMPLIFIER USING A CURRENT MATCHING ELEMENT.	9
FIGURE 2.1 - GBW OPTIMIZATION TECHNIQUES.	17
FIGURE 2.2 - INTEGRATION CONTOUR.	18
FIGURE 2.3 - TWO-TERMINAL INTERSTAGE DESIGN.	19
FIGURE 2.4 - CONSTANT K LOW-PASS FILTER.	20
FIGURE 2.5 - IMPEDANCE GIVING THE MAXIMUM FLAT GAIN: AMPLITUDE AND PHASE.	20
FIGURE 2.6 - FOUR-TERMINAL INTERSTAGE DESIGN.	22
FIGURE 2.7 - RC COUPLED CHAIN AMPLIFIER.	24
FIGURE 2.8 - CHAIN AMPLIFIERS: A) WITH TWO-TERMINAL INTERSTAGES; B) WITH FOUR-TERMINAL INTERSTAGES.	24
FIGURE 2.9 - BWER FUNCTIONS FOR TWO-TERMINAL (2T) AND FOUR-TERMINAL (4T) INTERSTAGES.	26
FIGURE 2.10 - STAGGER TUNED AMPLIFIERS: A) SINGLE TUNED; B) DOUBLE TUNED.	27
FIGURE 2.11 - 2 STAGES STAGGERED AMPLIFIER FREQUENCY RESPONSE.	30
FIGURE 2.12 - 2 STAGES STAGGERED AMPLIFIER, BWER AND OVERSHOOT (OV).	30
FIGURE 2.13 - DISTRIBUTED AMPLIFIER.	31
FIGURE 2.14 - COMPARISON OF AMPLIFYING STRATEGIES.	33
FIGURE 2.15 - INDUCTIVE PEAKING CIRCUITS: A) TWO POLE SERIES PEAKING, B) THREE POLE SERIES PEAKING, C) SHUNT-SERIES PEAKING, D) TWO POLE SHUNT PEAKING, E) THREE POLE SHUNT PEAKING.	36
FIGURE 2.16 - FREQUENCY RESPONSE UNDER MAXIMALLY FLAT CONSTRAINTS OF THE CIRCUITS OF FIGURE 2.15.	41
FIGURE 2.17 - BASIC T-COIL CIRCUIT.	41
FIGURE 2.18 - T-COIL INTERSTAGES: A) TWO-POLE T-COIL; B) THREE-POLE T-COIL; C) FOUR-POLE L+T.	42
FIGURE 2.19 - CAPACITIVE PEAKING CONCEPT.	44
FIGURE 2.20 - BWER USING CAPACITIVE PEAKING.	45
FIGURE 2.21 – ENHANCED CIRCUIT VERSION FIGURE 2.19.	45
FIGURE 2.22 - FEEDBACK AMPLIFIER.	46
FIGURE 2.23 - ROOT LOCUS TRAJECTORIES FOR MAXIMALLY FLAT FEEDBACK AMPLIFIERS HAVING 3 RD AND 4 TH ORDER DYNAMICS.	49
FIGURE 2.24 - FEEDBACK AMPLIFIER’S FREQUENCY RESPONSE: CONSTANT FEEDBACK (LEFT SIDE), VARIABLE FEEDBACK (RIGHT SIDE).	50
FIGURE 2.25 - SHUNT-SHUNT CAPACITIVE FEEDBACK.	51
FIGURE 2.26 - CAPACITIVE FEEDBACK; A) COMPLETE CIRCUIT, B) MILLER SIMPLIFIED CIRCUIT.	52
FIGURE 2.27 - BWER FUNCTIONS, USING RHP ZEROS.	54
FIGURE 2.28 – COMMON X CONFIGURATION.	55
FIGURE 2.29 - CASCODE AMPLIFYING STAGE.	56
FIGURE 2.30 - CHERRY-HOOPER AMPLIFIER: A) AND B) TRADITIONAL CX CHAIN AMPLIFIER (SMALL SIGNAL AND	

DETAILED SMALL SIGNAL EQUIVALENTS RESPECTIVELY); C) AND D) CHERRY-HOOPER TAS-TIS STAGE (SAME).	57
FIGURE 2.31 - BWER AND GAIN OF THE OF THE CHERRY-HOOPER AMPLIFIER.	59
FIGURE 2.32 - F_T DOUBLER: A) FUNCTIONAL CONCEPT; B) DIFFERENTIAL MODE HALF CIRCUIT.	59
FIGURE 2.33 - CURRENT-MODE AMPLIFYING TOPOLOGIES: A) CURRENT AMPLIFIER; B) VOLTAGE AMPLIFIER; C) TRANSIMPEDANCE AMPLIFIER; D) TRANSADMITTANCE AMPLIFIER.	61
FIGURE 3.1 - REFERENCE FEEDBACK AMPLIFIER.	67
FIGURE 3.2 - FIRST ORDER FEEDBACK AMPLIFIER.	68
FIGURE 3.3 - CLOSED-LOOP MAGNITUDE RESPONSE ADDING A DELAY ELEMENT TO THE FEEDBACK LOOP (PLUS DETAILED ZOOM OF THE CUT-OFF REGION).	69
FIGURE 3.4 - BWER AS FUNCTION OF THE LOOP-GAIN FOR A FIXED POLE-DELAY PRODUCT ($\delta=0,01$).	71
FIGURE 3.5 - CLOSED-LOOP FREQUENCY RESPONSE COMPARISON FOR MAXIMUM BWER AND MAXIMUM FLAT CONDITIONS.	72
FIGURE 3.6 - CLOSED-LOOP FREQUENCY RESPONSE WITH MULTIPLE PEAKS.	72
FIGURE 3.7 - REFERENCE AMPLIFIER WITH UNIT FEEDBACK.	74
FIGURE 3.8 - 2 ND ORDER DYNAMICS BWER CURVES.	77
FIGURE 3.9 - DELAYED FEEDBACK AMPLIFIER CLASSES: A) PURE DELAY; B) ALL-PASS DELAY; AND C) MIXED DELAY.	87
FIGURE 3.10 - MAXIMUM FLATNESS AND STABILITY CONSTRAINTS FOR A 1 ST ORDER DELAYED FEEDBACK AMPLIFIER.	88
FIGURE 3.11 - MAXIMUM FLATNESS AND STABILITY CONSTRAINTS FOR A 2 ND ORDER DELAYED FEEDBACK AMPLIFIER.	90
FIGURE 3.12 - MAXIMUM FLATNESS CONSTRAINTS FOR A 1 ST ORDER DELAYED FEEDBACK AMPLIFIER WITH MIXED DELAYS.	94
FIGURE 3.13 - ROOT-LOCUS FOR MIXED DELAY AMPLIFIERS FOR SEVERAL VALUES OF ϕ	95
FIGURE 3.14 - GAIN AND PHASE MARGIN FOR A 1 ST ORDER DELAYED FEEDBACK AMPLIFIER WITH MIXED DELAYS.	99
FIGURE 3.15 - GAIN AND PHASE MARGIN FOR A 2 ND ORDER DELAYED FEEDBACK AMPLIFIER WITH PURE DELAY.	100
FIGURE 3.16 - GAIN AND PHASE MARGIN FOR A 2 ND ORDER DELAYED FEEDBACK AMPLIFIER WITH MIXED DELAYS, FOR $\xi=50$	100
FIGURE 3.17 - 1 ST ORDER BWER GAIN CURVES FOR PURE AND ALL-PASS DELAY AMPLIFIERS.	102
FIGURE 3.18 - 1 ST ORDER BWER GAIN CURVES: EFFECT OF INTRINSIC DELAY.	103
FIGURE 3.19 - 2 ND ORDER BWER GAIN CURVES FOR PURE AND ALL-PASS DELAY AMPLIFIERS.	103
FIGURE 3.20 - 2 ND ORDER BWER GAIN CURVES: EFFECT OF INTRINSIC DELAY.	104
FIGURE 3.21 - 2 ND ORDER BWER GAIN CURVES: USING THE POLE SEPARATION.	104
FIGURE 3.22 - GRAPHICAL METHOD FOR BANDWIDTH ESTIMATION.	106
FIGURE 3.23 - GRAPHICAL METHOD TO FIND δ IN A 2 ND ORDER DELAYED FEEDBACK AMPLIFIER.	107
FIGURE 4.1 - GENERAL N PORT NETWORK.	113
FIGURE 4.2 - N-NETWORK WITH ELEMENT γ BETWEEN NODES U , V AND REFERENCE.	114
FIGURE 4.3 - CLASSICAL β -A MODEL.	115
FIGURE 4.4 - GENERAL FEEDBACK AMPLIFIER.	115
FIGURE 4.5 - G_∞ MEASUREMENT.	117
FIGURE 4.6 - VERTICAL BIPOLAR TRANSISTOR.	119
FIGURE 4.7 - NORMALIZED COMMON BASE CURRENT GAIN FOR DIFFERENT BASE DOPING PROFILES.	121
FIGURE 4.8 - NORMALIZED COMMON BASE SHORT-CIRCUIT CURRENT GAIN, COMPLETE AND APPROXIMATED MODELS FOR	

$\eta=0$; MAGNITUDE AND PHASE.	122
FIGURE 4.9 – GAAS HBT: CONSTRUCTION DETAIL.	123
FIGURE 4.10 – FIELD EFFECT TRANSISTOR, CONSTRUCTION DETAILS: A) JFET; B) MOSFET.	124
FIGURE 4.11 – HIGH FREQUENCY GAAS FET TRANSISTORS: A) MESFET, B) CONVENTIONAL HEMT, C) P-HEMT.	125
FIGURE 4.12 – FET SMALL SIGNAL MODEL COMPRISING DELAY.	126
FIGURE 4.13 – GENERAL Y PARAMETER REPRESENTATION OF THE DELAY ELEMENT.	127
FIGURE 4.14 – TRANSMISSION LINE R-L-C-G DISTRIBUTED MODEL.	128
FIGURE 4.15 – REFERENCE TRANSMISSION LINE SET-UP.	131
FIGURE 4.16 – FREQUENCY DEPENDENCE OF THE DELAY ACCORDING TO (4.39).	133
FIGURE 4.17 – MAGNITUDE OF THE FREQUENCY RESPONSE OF THE VOLTAGE TRANSFER FUNCTION (4.38).	133
FIGURE 4.18 – PLANAR TWO CONDUCTOR TRANSMISSION LINES: A) MICROSTRIP; B) STRIPLINE.	134
FIGURE 4.19 – MEANDER LINE.	136
FIGURE 4.20 – MEANDER LINE DELAY DESIGN EXAMPLE: EFFECTIVE DELAY (L) AND INPUT REFLECTION COEFFICIENT (S_{11}).	137
FIGURE 4.21 – AREA FILLING FACTOR.	138
FIGURE 4.22 – CONTOUR PLOTS OF THE AMPLITUDE ERROR E_A	142
FIGURE 4.23 – CONTOUR PLOTS OF THE PHASE ERROR, E_ϕ	142
FIGURE 4.24 – DEGENERATED COMMON X STAGE.	142
FIGURE 4.25 – DEGENERATED COMMON X STAGE, WITH DEGENERATING ELEMENTS.	144
FIGURE 4.26 – TRANSIMPEDANCE AMPLIFIER SMALL SIGNAL EQUIVALENT.	149
FIGURE 4.27 – DETAILED SMALL SIGNAL EQUIVALENT OF THE CIRCUIT OF FIGURE 4.26.	150
FIGURE 4.28 – CIRCUIT MODIFICATIONS TO EVALUATE: A) RETURN RATIO; B) DIRECT GAIN; C) ASYMPTOTIC GAIN.	151
FIGURE 4.29 – SMALL SIGNAL INPUT EQUIVALENT CIRCUIT OF THE AMPLIFIER IN FIGURE 4.26.	153
FIGURE 5.1 – CURRENT (A) AND VOLTAGE (B) AMPLIFICATION, USING VF AND CF.	162
FIGURE 5.2 – TRANSIMPEDANCE (A) AND TRANSADMITTANCE (B) AMPLIFICATION, USING CF AND VF.	165
FIGURE 5.3 – REFERENCE TRANSIMPEDANCE AMPLIFIER.	167
FIGURE 5.4 – TIA PLUS CMD DESIGN.	168
FIGURE 5.5 – NORMALIZED CUTOFF FREQUENCY AS A FUNCTION OF $(\omega_0 R_{IC} C_P)^{-1}$	170
FIGURE 5.6 – NORMALIZED FREQUENCY RESPONSE, EXPLORING PEAKING OUTSIDE THE BAND.	171
FIGURE 5.7 – GLOBAL FEEDBACK TIA PLUS CMD DESIGN.	172
FIGURE 5.8 – GENERIC THREE-PORT CC REPRESENTATION.	174
FIGURE 5.9 – CC-II SMALL SIGNAL EQUIVALENT CIRCUIT.	176
FIGURE 5.10 – SUPER-TRANSISTOR.	177
FIGURE 5.11 – CURRENT MEASURING USING: A) CC-I, B) CC-II, C) CC-III.	177
FIGURE 5.12 – CC-I AND CC-III IMPLEMENTATIONS, USING THE C-II; A) CC-I, B) AND C) CC-III.	178
FIGURE 5.13 – DIFFERENTIAL SIGNAL HANDLING WITH CC-II.	178
FIGURE 5.14 – CLASS A CC-II CIRCUIT TOPOLOGIES: A) SIMPLE ONE TRANSISTOR POSITIVE GAIN CC-II; B) NEGATIVE GAIN CC-II; C) IMPROVED POSITIVE GAIN CC-II.	180
FIGURE 5.15 – CC-II TOPOLOGIES BASED ON THE DIFFERENTIAL PAIR: A) NEGATIVE GAIN; B) POSITIVE GAIN.	181
FIGURE 5.16 – CLASS AB CC-II STRUCTURES: A) POSITIVE GAIN; B) NEGATIVE GAIN.	181

FIGURE 5.17 – LARGE GAIN CC-II: DESIGN CONCEPT.....	182
FIGURE 5.18 – CLASS AB CC-II INPUT STAGE.	183
FIGURE 5.19 – NORMALIZED CURRENTS I_N AND I_P (MATCHED TRANSCONDUCTANCE CASE, $R=I$, ON THE LEFT; CHANGING R FOR SEVERAL VALUES BETWEEN 0.4 AND 2.5, ON THE RIGHT).	185
FIGURE 5.20 – DYNAMIC RANGE DEPENDENCE ON R	186
FIGURE 5.21 – X, Y AND Z BRANCHES OF A CC-II: A) Y BRANCH WITH SIMPLE CURRENT SOURCE (CS); B) X BRANCH WITH MIRRORING TRANSISTORS (CM); AND C) Z BRANCH WITH CS.	187
FIGURE 5.22 – CLASS AB POSITIVE CC-II HALF CIRCUITS.	188
FIGURE 5.23 – POSITIVE CLASS A CC-II N HALF SMALL SIGNAL EQUIVALENT CIRCUIT: A) SIMPLIFIED VERSION; B) COMPLETE VERSION.....	190
FIGURE 5.24 - η_c AT DIFFERENT TECHNOLOGICAL NODES AND DIFFERENT FOUNDRIES.	193
FIGURE 5.25 – CC-II’S VOLTAGE GAIN AND CURRENT GAIN FREQUENCY RESPONSE (MAGNITUDE AND PHASE).....	195
FIGURE 5.26 – MOBILITY RATIO FOR DIFFERENT TECHNOLOGICAL NODES AND DIFFERENT FOUNDRIES.	197
FIGURE 5.27 – TERMINAL DRIVING POINT IMPEDANCES (ADMITTANCES); BOTH MAGNITUDE AND PHASE.....	198
FIGURE 5.28 – NON-IDEAL TRANSFER FUNCTIONS (REVERSE GAINS AND REVERSE TRANSFER ADMITTANCES); MAGNITUDE ONLY.	198
FIGURE 5.29 – CC-II’S BANDWIDTH PERFORMANCE EVALUATION.....	199
FIGURE 5.30 – FEEDBACK CONFIGURATIONS TO REDUCE THE X INPUT IMPEDANCE.....	202
FIGURE 5.31 – POSITIVE FEEDBACK SCHEME: SMALL SIGNAL EQUIVALENT CIRCUIT.....	203
FIGURE 5.32 – FEEDBACK EFFECT ON INPUT IMPEDANCE AND CURRENT GAIN, POSITIVE FEEDBACK SCHEME.....	205
FIGURE 5.33 – NEGATIVE FEEDBACK WITH OA SCHEME: SMALL SIGNAL EQUIVALENT CIRCUIT.	206
FIGURE 5.34 - FEEDBACK EFFECT ON INPUT IMPEDANCE AND CURRENT GAIN, NEGATIVE FEEDBACK WITH OA.....	207
FIGURE 5.35 - NEGATIVE FEEDBACK WITH TWO CC-IIS SCHEME: SMALL SIGNAL EQUIVALENT CIRCUIT.	208
FIGURE 5.36 - FEEDBACK EFFECT ON INPUT IMPEDANCE AND CURRENT GAIN, NEGATIVE FEEDBACK WITH TWO CC-IIS.	209
FIGURE 5.37 – PERFORMANCE COMPARISON.....	210
FIGURE 5.38 – DUAL LOOP POSITIVE-NEGATIVE FEEDBACK.	211
FIGURE 5.39 – DUAL LOOP POSITIVE-NEGATIVE FEEDBACK: EFFECT OF FEEDBACK ON IMPEDANCE.	213
FIGURE 6.1 – DELAYED FEEDBACK AMPLIFIER: A) EXPERIMENTAL SET-UP AND B) BUFFER CIRCUITRY.	217
FIGURE 6.2 – BWER EXPERIMENTAL RESULTS.	220
FIGURE 6.3 – FREQUENCY RESPONSE: MAXIMUM BWER OBSERVED UNDER MAXIMUM FLAT CONDITION.	220
FIGURE 6.4 – BOUNDED FREQUENCY RESPONSE NATURE.	221
FIGURE 6.5 – MAGNITUDE OF THE FREQUENCY RESPONSE: OSCILLATION CONDITION.....	221
FIGURE 6.6 - TIA CONCEPTUAL DESIGN.	222
FIGURE 6.7 – TRANSIMPEDANCE AMPLIFIER CIRCUIT.....	224
FIGURE 6.8 – TIA LAYOUT.....	225
FIGURE 6.9 – SIMULATED DELAY ELEMENT CIRCUIT.....	225
FIGURE 6.10 - PHASE AND AMPLITUDE ERROR WITH FIXED $C_{PX}=100FF$	226
FIGURE 6.11 - PHASE AND AMPLITUDE ERROR WITH FIXED $C_{PY}=100FF$	226
FIGURE 6.12 - PHASE AND AMPLITUDE ERROR WITH FIXED $C_{PZ}=100FF$	226
FIGURE 6.13 - MEASURED AND COMPUTED TIME DELAY.....	227

FIGURE 6.14 – TRANSIMPEDANCE AMPLIFIER: A) EXTERNAL SOURCE MODEL REPRESENTATION; B) BIAS MODEL.	228
FIGURE 6.15 – TIA SMALL SIGNAL MEASUREMENTS: FROM TOP TO BOTTOM, DIRECT GAIN, RETURN RATIO AND CLOSED- LOOP GAIN (MAGNITUDE AND PHASE).	229
FIGURE 6.16 – BWER AND RETURN RATIO THEORETICAL RESULTS VERSUS SIMULATED RESULTS.	231
FIGURE 6.17 – CUT-OFF FREQUENCY AND OPTIMIZING VALUES OF C_Z	232
FIGURE 6.18 – NOISE FREQUENCY RESPONSE.	233
FIGURE 6.19 – TIA PLUS CMD SIMULATION SET-UP.	233
FIGURE 6.20 – CLASS AB TRANSLINEAR CC-II.	235
FIGURE 6.21 – CC-II OPTIMIZATION SET-UP.	235
FIGURE 6.22 – γ OPTIMIZATION.	236
FIGURE 6.23 – CURRENT AND VOLTAGE GAIN FREQUENCY RESPONSE FOR SEVERAL VALUES OF W_N	237
FIGURE 6.24 – TERMINAL DRIVING POINT IMPEDANCES (ADMITTANCES) FOR SEVERAL VALUES OF W_N	238
FIGURE 6.25 – NON-IDEAL TRANSFER FUNCTIONS (REVERSE GAINS AND REVERSE TRANSFER ADMITTANCES).	238
FIGURE 6.26 – EFFECT OF W_N ON CURRENT GAIN AND X INPUT IMPEDANCE.	239
FIGURE 6.27 – EFFECT OF I_{BIAS} ON CURRENT GAIN AND X INPUT IMPEDANCE.	240
FIGURE 6.28 – EFFECT OF I_{BIAS} AND W_N ON NOISE FLOOR.	240
FIGURE 6.29 – CC-II LAYOUT.	241
FIGURE 6.30 – CMD SIMULATION SET-UP.	241
FIGURE 6.31 – X INPUT IMPEDANCE REDUCTION.	243
FIGURE 6.32 – IMPEDANCE CONTROL AND BWER FUNCTIONS.	243
FIGURE 6.33 – TIA+CMD SIMULATION SET-UP.	244
FIGURE 6.34 – BANDWIDTH AND BWER BEHAVIOR FOR SEVERAL VALUES OF C_p	246
FIGURE 6.35 – EQUIVALENT NOISE INPUT CURRENT FREQUENCY BEHAVIOR.	246

Acronyms

AMS	Austria Microsystems
APD	Avalanche Photo Detector
BC	Base Collector Junction
BE	Base Emitter Junction
BiCMOS	Bipolar and CMOS process
BJT	Bipolar Junction Transistor
BWER	Bandwidth Enhancement Ratio
CC	Current Conveyor
CC-I	First Generation Current Conveyor
CC-II	Second Generation Current Conveyor
CC-III	Third Generation Current Conveyor
CCCS	Current Controlled Current Source
CF	Current Follower
CM	Current Mirror
CS	Current Source
CX	Common X
CY	Common Y
CZ	Common Z
CMD	Current-Matching Device
CML	Current Mode Logic
CMOS	Complementary Metal-Oxide-Semiconductor
DC	Direct Current
DLL	Delay Locked Loop
DUT	Device Under Test
ECL	Emitter Coupled Logic
FDE	Functional Differential Equation
FET	Field Effect Transistor
FO	Fiber Optic
FSO	Free Space Optical
GBW	Gain-Bandwidth product
HB-LED	High Brightness LED
HBT	Hetero Junction Bipolar Transistor
HEMT	High Electron Mobility Transistor
IC	Integrated Circuit
JFET	Junction FET
LED	Light Emitting Diode
LHP	Left Half Plane
LOS	Line of Sight
MESFET	Metal-Semiconductor FET
MOSFET	Metal-Oxide-Semiconductor FET
NIC	Negative Impedance Converter
NPN	N type – P type – N type BJT
PCB	Printed Circuit Board
PD	Photo Detector
P-HEMT	Pseudomorphic HEMT
PIN	P type – Intrinsic – N type diode
PNP	P type – N type – P type BJT
RHP	Right Half Plane
TAS-TIS	TransAdmittance Stage - TransImpedance Stage
TIA	Transimpedance Amplifier
TEM	Transverse Electrical Magnetic

TL	Transmission Line
OA	Operational Amplifier
OTA	Operational Transconductance Amplifier
UV	Ultra Violet
VCCS	Voltage Controlled Current Source
VCVS	Voltage Controlled Voltage Source
VF	Voltage Follower
VOA	Voltage Operational Amplifier

List of Frequent Symbols:

α	Normalized frequency variable, attenuation constant and transistor CY current gain	f_o	Wheeler index frequency
α_c, α_i	normalized cut-off and limit frequencies	f_T	Transition frequency
ξ	Damping Coefficient	f_α, f_β	α and β cut-off frequencies
δ	Pole-delay product	f_{max}	Maximum oscillating frequency
$\delta(s)$	Characteristic quasi-polynomial	$G(s)$	Loop-Gain transfer function
β	Feedback factor, phase constant and transistor CX current gain	G_∞	Asymptotic gain
τ	General time constant and intrinsic delay	G_o	Direct gain term
τ_B, τ_C	Base and collector transit times	G_o	Loop-gain and loop-gain transfer function
γ	Propagation constant	G_x	General conductance
ω	Real frequency variable	g_m	Transistor's transconductance
ω_n	Natural Frequency	H	Closed-loop transfer function
$\omega_c, \omega_r, \omega_l, \omega_m$	cut-off, reference cut-off, limit and maximum cut-off frequencies	H_L, H_U	Closed-loop bounding functions
ω_g, ω_p	Gain and phase cross-over frequencies	I_x, i_x	current at node x
$\Delta(s)$	Padé delay transfer function	K	Bandwidth Enhancement ratio function
Δ, Δ_{ij}	Admittance matrix and cofactor	k_{ij}	Generalized k-port parameters
Δ_G, Δ_ϕ	Gain and phase margins	L, L_{eff}	Complex exponential delay
Δ_T	Delay error	L	FET length
η_o	Characteristic impedance of the free space	L_D	Diffusion Length
μ_o, μ_r	Open-air and relative permeability	$N(), n_r(), n_i()$	Numerator polynomials, real and imaginary parts
ϵ_o, ϵ_r	Open-air and relative permittivity	p_x	Pole angular frequency
ρ_x	Reflection coefficient	R	Return ratio
A_x, A	General voltage transfer functions	R_r	Null return ratio
a_k	denominator time constants	R_x	General resistance
B	Bandwidth	r_π, r_e, r_i	Transistor's intrinsic resistances
b_k	Numerator time constants	s	Complex frequency variable
c	Velocity of light	T	Padé approximated delay
C_x	General capacitance	V_x, v_x	Voltage at node x
c_π, c_x, c_μ, c_z	Transistor's intrinsic capacitances	V_t	Thermal gap voltage
$D(s)$	General delay transfer function	V_T	FET threshold voltage
$D(), d_i(), d_r()$	Denominator polynomials, real and imaginary parts	y_{ij}	y-port parameters
d	Pole separation factor	v_p	Wave propagation velocity
E	Electric Field	Y_x	General admittance
E_ϕ	Phase error	W	FET width
E_A	Amplitude error	W_B	Base width
F	Return difference	w, l, h, s, t	General purpose width, length, height, separation and thickness measures
f	Linear frequency	Z_{ib}, Z_{ij}, Z_x	Driving point, transfer and general purpose impedances
		Z_o	Characteristic impedance
		z_x	Zero angular frequency and RHP zero time delay

1 INTRODUCTION

"The composition of vast books is a laborious and impoverishing extravagance. To go on for five hundred pages developing an idea whose perfect oral exposition is possible in a few minutes! A better course of procedure is to pretend that these books already exist, and then to offer a resume, a commentary . . . More reasonable, more inept, more indolent, I have preferred to write notes upon imaginary books." - Jorge Luis Borges

Summary

Achieving large gain and bandwidth in amplifiers is an old problem. New solutions to this problem remain a current theme of investigation. In this thesis, the design of high gain and bandwidth transimpedance amplifiers suitable for FSO communication receivers is addressed. This chapter introduces the nature of GBW limitations in these FSO front-ends. Some known methods to combat GBW in amplifiers are briefly presented, followed by a general analysis of the proposed solutions. The thesis development framework is also discussed, highlighting past developments and future perspectives for FSO systems. Finally, the original contributions of this work and the thesis organization are presented.

1.1 Problem Statement

Maximizing both the gain and bandwidth of an amplifier is an old and largely debated problem in electronic circuit design. Since the early days of electronics that the need for amplifiers having high gains and large bandwidths has been a major force motivating the search for new and better solutions. It was during the electronic valve era that this problem first acquired its renown importance [247, 248]. Since thermionic valves had limited frequency response, it was of paramount importance to designers to know in advance the gain and frequency requirements for his amplifiers. Bode was the first to formalize the problem and to deliver a solution [248], revealing that the product of an amplifier's gain times its bandwidth (the gain-bandwidth product, GBW) is generally fixed by the ratio g_m/C , where g_m represents the transconductance of the active devices within the amplifier and C a combination of parasitic capacitances at the ports of the amplifier. An immediate consequence of Bode's GBW limits is that, gain and bandwidth are reciprocals, one can not be increased without sacrificing the other. Transistor and integrated circuit technologies reduced this problem. The newly developed active devices had enhanced gain capabilities. Parasitic capacitances arising from both wiring issues and active devices were also reduced, allowing larger GBWs to be achieved. Nevertheless, the ratio g_m/C remained the fundamental GBW limitation, simply this limit was not critical for the majority of applications.

Nowadays GBW limits play again a key role in the design of amplifiers. There are a wide range of applications demanding higher transmission rates and higher sensitivities, thus requiring large gain and bandwidth amplifiers. One particular case on which these requirements are often difficult to meet is the design of amplifying front-ends for free-space optical (FSO) receiving systems.

1.1.1 Free-Space Optical Front-Ends

Figure 1.1 depicts a simplified diagram of the physical layer of a FSO communication system [270, 271, 279, 284]. This diagram consists of three main components: the emitter, the channel and the receiver. The emitter process and transforms the input data. To this end, a pulse shaping stage

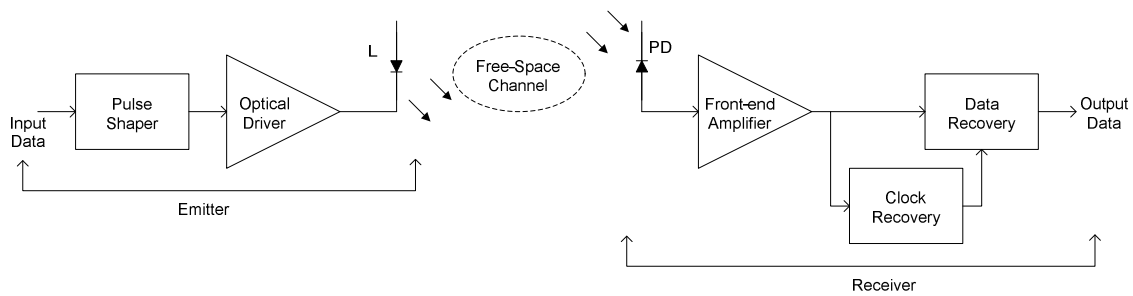


Figure 1.1 - FSO communication system (physical level).

adapts the input data to an efficient line code suitable for transmission. The optical driver converts the electrical impulses from the pulse shaper into the optical domain, using a light emitting device (L in Figure 1.1), which can cover the complete visible light wave length range or just (more commonly) a particular range (as is the case of infrared [1, 4, 8] or ultraviolet [3] communication based systems). The optical signal travels on the free-space between the emitter and receiver, suffering the influence of multiple noise sources (ambient illumination, in particular, is the dominant noise source), and channel attenuation. The receiver detects, amplifies and recovers the optical signals reaching its photo-detector (PD in Figure 1.1). The photo-detector converts the optical signal into an electrical current. Usually, PIN (P type – Intrinsic – N type) type photo-diodes or photo-transistors are the selected choice for this purpose. The converted current is generally very small, requiring a front-end amplifier to amplify the signal to adequate levels before further processing. Finally, the clock and data recovery units recover and synchronize the detected data.

The design of front-end amplifiers adequate for these FSO receivers is particularly difficult to accomplish. This kind of amplifiers pose several design challenges [293], namely: i) high sensitivity, thus demanding high-gain stable configurations and low input referred noise; ii) the constant demand of higher transmission rates implies increasingly larger amplifier bandwidths; iii) high electromagnetic immunity (which can be achieved using differential signal path topologies); and iv) high input dynamic range (eventually achieved using switched or variable gain topologies). Design the front-end amplifier in order to meet the first two requirements, high-gain and high-bandwidth, is perhaps the most challenging problem. The other design considerations are usually simpler to achieve, usually without seriously impairing nor gain neither bandwidth.

The main reason for GBW limitation on these FSO amplifiers and the reason why these two requirements are so difficult to achieve, comes from the input circuitry. Figure 1.2 shows a simplified small signal equivalent of the front-end input circuitry. In order to maximize the optical signal detection, these optical receivers must use photo-detectors with large photo-sensible areas,

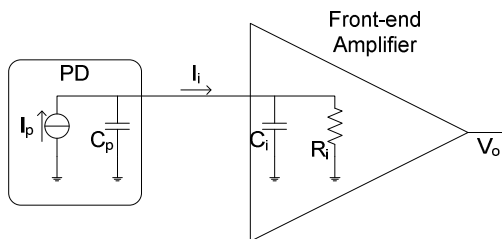


Figure 1.2 – Small signal equivalent of the input circuitry of the front-end amplifier.

which result in large intrinsic capacitances. From a circuit design perspective this is adequately represented by a current source, I_p , (representing the converted photo-current), in parallel with the intrinsic capacitance, C_p , as depicted on Figure 1.2. The amplifier input impedance is generally a complex function. However, for illustration purposes it can be assumed as a first order parallel

association of a resistance R_i and a capacitance C_i . Using this simplified model, the pole contribution due to the input circuitry becomes ruled by $R_i(C_p+C_i)$. Assuming that it is possible to

design the amplifier in such a way that the other poles in the system have smaller associated time constants, then the input circuit time constant is the dominant one. Thus, in order to meet the high-bandwidth requirements, the input circuit time constant should be minimized¹. There are two possibilities to minimize the input time constant: i) use better photo-detectors with smaller intrinsic capacitances – in general this implies increasing the overall budget of the project or reducing the active area of the photo-detector (which, in general impair system performance [6, 7, 9, 12]); ii) reduce the front-end input impedance. The second method is always preferred since it does not imply drastic changes at the system level nor circuit performance.

A common strategy to reduce input impedance in amplifiers is to use adequate feedback configurations. Transimpedance amplifiers are in this sense the most adequate. First of all, transimpedance amplifiers exhibit a desired current to voltage transfer characteristic. Transimpedance amplifiers can be constructed using feedback configurations, as is the case of shunt-shunt feedback represented on Figure 1.3. Transimpedance amplifiers represent the best compromise between gain and bandwidth². Assuming the simplified model of Figure 1.3, the input time constant is approximately given by $R_{if}(C_i + C_p)$, where R_{if} represents the input resistance with feedback. At a first sight, it seems that the input impedance can be made as smaller as required, thus reducing the effect of

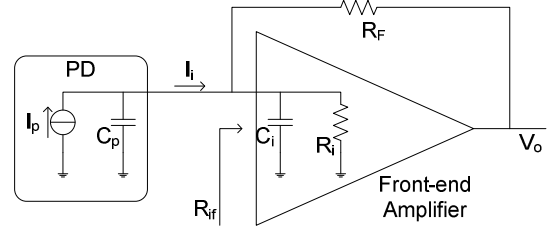


Figure 1.3 – Transimpedance front-end.

the input time constant. However, a close inspection shows that this is not as straightforward as it seems. For a reasonable well design transimpedance amplifier, the gain with feedback should be as close as possible to the value of the feedback resistance, R_F . On the other hand, a meaningful approximation of R_{if} is given by R_F divided by the voltage gain (assuming that loading effects posed by input and output circuits are negligible and the voltage gain is large when compared to unity). Both gain and input impedance are direct related through R_F , so in general reducing the input impedance implies reducing the transimpedance gain. It is possible to increase voltage gain in order to reduce R_{if} . A simple strategy to increase voltage gain is to include more gain stages on the forward amplifier. This allows larger gains to be achieved but implies an increased complexity of

¹ In general the amplifier's input impedance is more complicated than this simple model. Nevertheless the association of the photo-detector intrinsic capacitance together with resistive part of the input impedance act as a major bandwidth restriction factor.

² Other alternatives resort to high input impedance amplifiers (with large gain and restricted bandwidth) or small input impedance amplifiers (with small gain and large bandwidth) [270, 271].

the system dynamics. As the number of stages is increased, the pole-zero contributions due to each added stage becomes rather evolved, posing serious restrictions on the closed-loop stability. FSO amplifiers thus have serious GBW challenges.

1.2 Methods to Combat GBW Limitations

The GBW problem in amplifier design is not exclusive to feedback transimpedance amplifiers. Since Bode's results that the nature of GBW limitations has been acknowledge as a fundamental issue and thus intrinsic to all possible amplifier configurations [248]. The work of Bode on GBW limitations has served as the grounding base for the development of efficient methods to approach and overcome these limitations. Some of these methods are also suitable to the design of transimpedance amplifiers, with and without feedback. For instance, distributed amplification (original discovered in 1937 by Percival [42] and generalized in 1948 by Ginzton [43]) appears to be an amplifier design technique not restricted in Bode's sense. Using the concept of distributed amplification it is possible to explore gain over frequency ranges above Bode's limits. Distributed amplifiers have been used with success to conceive broadband amplifiers suitable for fiber optic receivers and high frequency measurement equipment [266]. Other approaches to increase gain without severe cuts on bandwidth resort to chain amplification or even paralleled chain amplification (as suggested by Linvill in 1950 [30]). Chain amplification allows gain to be distributed over large or different bandwidths, thus reducing bandwidth penalties. In this sense, the concept of stagger-tuning, invented in 1948 by Wallman [34, 35], promote a optimum strategy to distribute bandwidth in chain amplifiers.

Current-mode concepts are also well suited for high-bandwidth and moderate gain achievements [259, 263, 276]. The philosophy behind current-mode design is to design circuits able to process signals on the current domain. Traditional voltage-mode circuits process signals as node voltages. Since the medium inside a integrated circuits is predominantly capacitive, large swing node voltages imply large amounts of time wasted to charge/discharge parasitic capacitances. These bandwidth limitations can be efficiently reduced if signals are taken as currents flowing on circuit branches. This line of reasoning is particularly accurate for circuits consisting only on transistors, since for this case large current swings correspond to compressed voltage swings due to the nature of the transistor's characteristics (logarithmic compression for bipolar transistors and square root compression for field effect devices). High-bandwidth amplification using current-mode concepts is based on the existence of two elements: voltage and current followers. These are amplifier configurations having the largest possible bandwidth. Gain relations are then obtained by the simple combination of these elements and pure resistances [276].

The aforementioned techniques are thus suited to avoid and even surpass the Bode limits, but imply dramatic changes at the circuit design level. More traditional techniques try to explore the maximum possible GBW for a prescribed configuration; these are known as frequency optimization techniques[286]. Two approaches for frequency optimization are inductive peaking [48-63] and capacitive peaking [64-74]. Both of these techniques try to explore the presence of complex poles within the dynamics of the amplifier as a means to achieve peaking effects on the frequency response. It is possible to extend the amplifier bandwidth if these peaking effects can be made to occur near to the amplifier's cut-off frequency.

Inductive peaking is more mature than capacitive peaking. The effects of inductive peaking were originally observed by filter designers during the early stages of the thermionic valves era [247]. Inductive peaking is based on resonance phenomena common to all RLC circuits. Amplifiers are predominantly active RC circuits, nevertheless, it is possible to include inductive elements in some special locations of a gain stage in order to promote resonance.

Capacitive peaking is rather evolved. Its first appearance was reported in 1968 by Mataya [64], as a means of compensation on transistor's frequency response. Its generalized use as a tool suitable to amplifier design appeared much later by Chien in 1999 [70]. In its basic essence, capacitive peaking is not different from inductive peaking. The optimization of frequency response is achieved through the exploitation of the same resonance phenomena. The unique difference is that capacitive peaking does not involve inductive elements to achieve resonance. Instead of this, capacitive peaking explores the presence of complex conjugated pairs of poles in the dynamics of the amplifier. This means that capacitive peaking is always associated to active circuits able to synthesize the same dynamics found in RLC circuits. It is not surprising to find that capacitive peaked circuits are generally feedback circuits [70].

The aforementioned methods are based on general concepts that can be applied to a wide range of amplifier configurations. There are also certain specialized amplifier configurations that offer high GBW capabilities. Among these cascode stages [122, 123], Cherry-Hopper amplifiers [125-131] and f_T doublers [131], are certainly between the most representative. Both of these amplifier configurations were advanced as alternatives to traditional single transistor configurations, producing the same gain with large bandwidth capabilities.

1.3 Proposed Solutions

This thesis presents two techniques suitable for the design of FSO front end amplifiers demanding high GBW with high capacitance photo detectors at the input. The first technique is a general method for bandwidth enhancement in feedback amplifiers. In its essence, this is another

form of frequency optimization, based on the exploitation of small delay elements inside the feedback loop of feedback amplifiers. The second technique finds application in the design of transimpedance amplifiers with a dominant input pole, as it is the case of FSO amplifiers. As it was previously discussed, the GBW limitations arise due to the interrelationships of transimpedance gain and input impedance often found in these amplifiers. One possible way to reduce this problem consists precisely in breaking this transimpedance gain versus input impedance dependence. This thesis proposes the usage of active current matching circuits able to promote the isolation between the photo detector and the transimpedance stage. This current matching concept allows the designer to specify transimpedance gain and input impedance in an independent fashion.

1.3.1 Delayed Feedback Concept

The concept of delayed feedback arises from the exploitation of small time-delay elements that coexist within the amplifier circuitry (as depicted in Figure 1.4). Naturally, the presence of these small delays in feedback amplifiers has been acknowledge several years ago. In 1940 Shaw issued a patent on feedback amplifiers comprising a delay element in its feedback loop [77], showing how that delay should be treated in order to avoid stability problems. By that time amplifiers comprising electronic valves had large physical dimensions, with large delays associated with the wiring between circuit's elements. Stability in such amplifiers became a serious issue as gain and bandwidth requirements increased. With the invention of the transistor and integrated technology, these problems were partially solved, since the delays became negligible as the circuit's dimensions shrunk. Nevertheless, the nature of the problem persisted and eventually resurfaced. Recently, Nowack reported the manifestation of the same sort of stability problems arising in high gain and bandwidth feedback transimpedance amplifiers suited for fiber optic systems [102]. Nevertheless, modern feedback design techniques still disregard delay effects [267, 277, 279]. The effect of delays in feedback amplifiers is however becoming of paramount importance, not only from the stability view point (where these small delays can compromise amplifier operation), but also from the frequency response view point (where these delays can be used as another means to promote peaking phenomena).

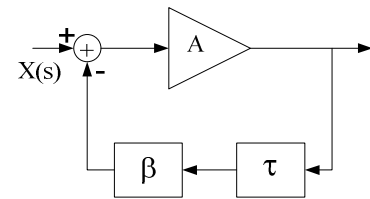


Figure 1.4 – Delayed feedback amplifier (conceptual diagram).

The concept of frequency peaking in feedback amplifiers can be easily explained. In a feedback amplifier having more than one pole in its internal dynamics, the position of the poles depends on the loop-gain of the amplifier. Starting with two poles placed on the negative real axis, as the loop-gain increases the two poles come close together and eventually depart from the real axis, thus

becoming complex. As the two poles get nearer, the bandwidth increases, since the dominant pole shifts to higher frequencies on the real axis. Increasing further the loop-gain, the two poles become complex conjugate and the bandwidth increases further, until the critical damping of the two poles is reached. Loop-gain above this critical point imply large overshoot on the frequency response and loss of bandwidth. Similar peaking phenomena is also verified for amplifiers having more than two poles. The mechanism that seems to justify these peaking phenomena is the possibility of having complex poles with variable damping coefficients. Similar effects can be explored if the amplifier's internal dynamics include one delay element. According to standard root-locus techniques, such amplifiers have an infinite number of poles, thus similar frequency peaking can be expected from such arrangement. In a delayed feedback amplifier (a feedback amplifier comprising one delay in its feedback loop – see Figure 1.4), the maximum possible bandwidth enhancement depends on loop-gain, delay and dominant set of poles. Theoretically this maximum has the value of $1 + \sqrt{2}$, meaning that the amplifier bandwidth can be enhanced by almost 140%, in relation to the standard non-delayed feedback case [296, 301].

On the framework of this thesis, the concept of delayed feedback was developed as a general frequency optimizing tool that can be applied to any feedback amplifier configuration. The presence of delay elements arising due to amplifier internal circuitry can be explored for optimization purposes. It is also possible to include inside the feedback loop a specific designed delay element in order to explore the same frequency optimizing mechanism without impairing stability.

1.3.2 Current Matching Concept

The previous discussion on FSO front-end amplifiers revealed that the intrinsic capacitance of the photo-detector is a major bandwidth limiting factor. These capacitances together with the amplifier's input impedance, pose a low frequency dominant pole that restricts bandwidth. Transimpedance configurations can partially address this issue. However, since in a transimpedance amplifier the input impedance depends strongly on the desired closed loop-gain, aiming for low input impedance also implies low gain. The required solution must break this gain-impedance dependency.

The proposed solution relies on the usage of an active current matching element placed between the photo-detector and the transimpedance stage, as depicted on Figure 1.5. This current matching element (CMD) provides the following improvements: i) isolation between the photodetector and the transimpedance stage – gain bandwidth achievements of the transimpedance amplifier are fulfilled independently from the photodetector capacitance; ii) low input impedance – required to solve the bandwidth restriction problem; and iii) an additional current gain factor – which naturally

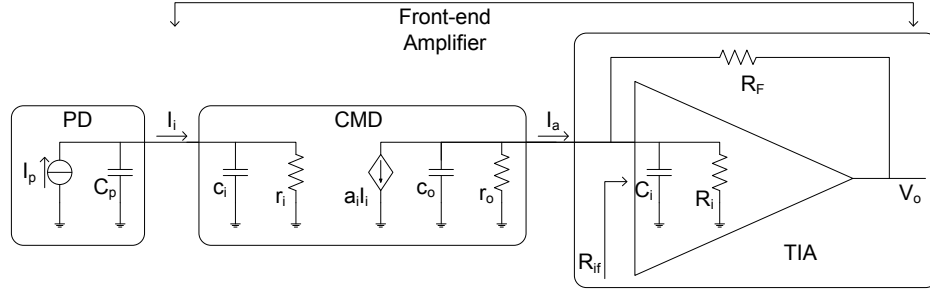


Figure 1.5 – FSO front-end amplifier using a current matching element.

improves the overall gain. The CMD should be able to convey a replica of the input current, from a low impedance input port to a high impedance output port. This suggests that the CMD is a current controlled current source, measuring the input current and providing an output replica. Current-mode design and particularly, current conveyors, are in this sense suitable candidates to synthesize the CMD stage.

1.4 Thesis Framework

Devising new GBW optimizing techniques suited to amplifier design is per se a sufficient and actual reason to advance a PhD study program. Joining to this, the expertise of the Circuits and Integrated Systems group of the Instituto de Telecomunicações – Aveiro, on the development of front-ends for free-space infrared communication receivers, provides a set of conditions to support and fulfill the requirements of this work (see for instance [12] and the references therein).

Being able to cope with both international market and academic trends was also an important motivating aspect. Current trends toward solid-state lighting technologies open new design perspectives on free-space optical communication systems [22]. The exploitation of visible light as a means of conveying digital or analog information between communication terminals is one such perspective, where previous infrared expertise is most welcomed.

1.4.1 Wireless Infrared FSO Systems

Wireless infrared communications systems were first introduced by Gfeller and Bapst in 1979 [1]. Their original work demonstrate the feasibility of infrared radiation to convey information signals in the free space optical domain. Since then, wireless infrared communication systems have been object of active research, by both academia and industrial sectors [4]. According to Kahn and Barry, in their 1997 seminal contribution on the topic [8], these kind of systems can be classified into six categories based on two orthogonal properties: the existence of line-of-sight (LOS) between emitter and receiver; and the directional properties of both emitter and receiver (which can be both directive, non-directive and hybrid). LOS configurations favors the existence of direct communication channel between emitter and receiver, while non-LOS configurations must rely on

reflection mechanisms that the optical infrared signal must undergo between the emitter and receiver.

The architectural aspects of the emitter and receiver are naturally of major importance for overall system design [8]. Several fundamental problems originated from these considerations, namely: channel modeling – occupied with the study of the degradation effects on the (infrared) signals, posed both by attenuation, reflection mechanisms and the existence of multiple noise sources (due to ambient illumination); emitter design – the choice between directed or non-directed emitter topologies impair the (infrared) signal distributions and hence, receiver design; receiver design – the study of the receiver characteristics able to cope with all the aforementioned problems, including sensitivity, bandwidth, noise impact and LOS or reflection based channel.

The Department of Electronics, Telecommunications and Informatics from the University of Aveiro developed a considerable expertise on all the aforementioned fields. Several works leading to master and doctoral degrees have been published since 1990. The most relevant that deserve mention are: i) Valadas which, originally proposed the infrared sectored receiver [6]. Such receivers are able to cope optimally with the stringent requirements posed by diffuse emitter-receiver configurations. In these receivers the optical signal can travel by several different paths between emitter and receiver, thus suffering from different optic interfering sources (reflections, noise, attenuation); ii) The work of Moreira [7], dealing with problems of interfering noise sources and their models; iii) The work of Lomba [2, 9], on simulation, modeling and optimization of the optic channel and circuit design issues; iv) The first attempt to design all the receiver using integrated circuit design was issued by Aguiar in 1994 [5]. The group of Integrated Circuits and Systems from Instituto de Telecomunicações, Aveiro section, also add its contribute to the design of integrated systems developing complete optical receivers [12, 14-16] and transimpedance front-ends having large GBWs [10, 11, 17].

1.4.2 Visible Light Communication Systems

Until now it was not feasible to establish communication links using the available illumination technology. The reason for this lies on the fact that traditional lamps have poor switching capabilities [22]. There are mainly two types of lighting devices involving electric currents, incandescent bulbs and discharge lamps. Incandescent bulbs consist of a metal filament that glows when heated by an electrical current. Such glowing phenomena, known as incandescence, has large inertia to variations on the electrical current; the metal wire cannot cool instantaneously. Discharge lamps use a different phenomenon, known as photoluminescence; the ability to convert UV (Ultra Violet) radiation of an ionized gas (mercury vapor) by a phosphor coating. As in the previous case,

the ionized gas cannot be instantaneously switched. With these limitations, the unique communication mechanism is restricted to primitive light signals in existence since ancient times.

The phenomenon of electroluminescence, observed for the first time in 1907 in a silicon carbide (carborundum, also known as galena crystal) semiconductor, became popular during the sixties with the invention of LED (light emitting devices) [22]. LEDs were certainly not substitutes to traditional lamps. Nevertheless, they opened the way to new illumination perspectives based on solid-state devices. Recent advances brought HB-LEDs (high brightness LEDs). These LEDs are made of III-V semiconductor compounds. These HB-LEDs can achieve the same illumination efficiencies as traditional lighting solutions, with reduced power and area occupancy, which makes them suitable for a wide range of environments. Nowadays their widespread application to all sort of indoor and outdoor environments is becoming evident. For instance, there are cases of entire buildings using solid-state lighting solutions both in the public (such as bridges) and private (residential buildings) domains [22]. Their usage in the automotive market is also under fast development: the substitution of car luminaries by solid-state HB-LEDs or even in traffic control [18, 22].

One important difference between these HB-LEDs and traditional electric lamps is their superior switching capabilities. HB-LEDs bear the promise of high frequency switching operation; HB-LEDs are presently developed for illumination purpose alone, nevertheless they can be used to transmit signals with frequencies up to 10MHz. International academia is starting to explore this new technology to devise new optical communication scenarios based on visible light [13, 19-21]. A promising example is the delivery of traffic information between traffic lights and circulating cars [18]. A whole set of new problems arise in these visible light communication scenarios, some of the most prevailing are: i) channel modeling; ii) sun light interference; iii) available detectors; and iv) transmitter-receiver design. Some of these problems benefit from experience of infrared technology, namely optical channel modeling and receiver design are quite similar. The currently available photodetectors are mainly existing photodiodes with short wavelength range, tuned for infrared systems. However, this present state is changing rapidly and better photodetectors, able to cope with large visible wavelength spectrum are expected to emerge in a near future.

1.5 Software Support

The work reported on this thesis was accomplished using several software developing tools available at the University of Aveiro and in the Telecommunications Institute. These tools included text editing software, mathematical simulation software and circuit design and simulation software. Text editing was accomplished using Microsoft based products running on a Windows XP based

platform, which included: Microsoft Word 2003, Microsoft Visio 2003, Microsoft Excel 2003 and Mathtype 5.2. Paint Shop Pro 9 was used for picture editing. Numerical simulation and analytic mathematical results were accomplished and verified using Matlab 6.5 and Mathematica 5.1 respectively. Circuit design used Cadence Design Framework II (distribution 2003/2004) for electronic circuit design, layout design and electronic simulation. Finally, electromagnetic simulation was accomplished with ADS 2005.

1.6 Original contributions

A total of fourteen paper contributions arising from this work were published, both in IEEE conference proceedings and internationally acknowledge journals (Kluwer's International Journal of Analog Circuits and Signal Processing and Wiley's Journal of Circuit Theory an Applications), namely:

- The concept of active matching between the photodetector and the transimpedance amplifier using current conveyors was published and presented in four conference contributions, both national and international, namely: ICECS 1999 - International Conference on Electronics Circuits and Systems 1999, realized in Paphos, Cyprus; ISCAS 2001 - International Conference on Electronics Circuits and Systems 2001, realized in St. Julians, Malta; DCIS 2001 - Design of Circuits and Systems 2001, realized in Porto, Portugal; and ConfTele 2003 - Conferência de Telecomunicações 2003, Aveiro, Portugal . An international journal paper also originated from this concept, published on the International Journal of Analog Circuits and Signal Processing 2002, Kluwer Academic Publishers [287-289, 292, 293].
- Two contributions on general analysis of current mirrors frequency response and noise behavior, both in ICECS 2002 - International Conference on Electronics Circuits and Systems 2002, realized in Dubrovnik, Croatia [290, 291].
- One contribution on symbolic circuit analysis methods, using Laplace expansion method to simplify the analysis flow of some classes of active RLC circuits, presented at ICECS 2003 - International Conference on Electronics Circuits and Systems 2003, realized in Sharjah, United Arab Emirates [294].
- A new technique to impedance reduction using current conveyors presented in ICECS 2004 - International Conference on Electronics Circuits and Systems 2004, realized in Tel-Aviv, Israel. This technique used a dual loop feedback scheme with both positive and negative feedback [295].
- The delayed feedback concept was published in four different international conference proceedings. Two contributions about the theoretical aspects of the concept published in

ECCTD 2005 - European Conference on Circuit Theory and Design 2005, Cork, Ireland, and ISCAS 2005 - International Symposium on Circuits and Systems 2005, Kobe, Japan. Two contributions about circuit design procedures, in ICECS 2005 - International Conference on Electronics Circuits and Systems 2005, realized in Ganmarth, Tunisia, and ISCAS 2006 - International Symposium on Circuits and Systems 2006, Kos, Greece. One further contribution concerning this concept was published on Wiley's International Journal of Circuit Theory and Applications 2007 [296- 299, 301].

- One contribution on analog delay cell design using Pade approximants, published in ICECS 2006 International Conference on Electronics Circuits and Systems 2006, realized in Nice, France [299].

1.7 Chapter Overview

The work reported in this thesis is divided in seven chapters. This chapter discussed motivation and overview aspects. The content of the following chapters is organized as follows:

- Chapter 2 describes the state-of the art. A revision of the fundamental GBW limitations based on Bode's work is presented. Structures and techniques used to approach or surpass these limitations are also briefly discussed.
- Chapter 3 presents and explores the concept of delayed feedback. The basis for flat gain operation is addressed in this chapter, posing restrictions on loop-gain, delay, and circuit time constants. The general design methodologies are based on design curves that relate both delay, loop-gain and time constants, allowing a method to predict the bandwidth improvement.
- Chapter 4 presents circuit design methodology focused on the proposed delayed feedback concept. General return ratio analysis techniques are proposed to the analysis of delayed feedback amplifiers. Delay element analysis and synthesis methods are explored focusing on both intrinsic delays and designed delays. Finally, the complete optimization procedure is applied to the analysis of a generic transimpedance amplifier.
- Chapter 5 presents the proposed current matching device design methodology. Several current conveyor based topologies able to produce controllable, low input impedance CMD are discussed. Since these devices are in general feedback systems, the integration of the delayed feedback concept is also briefly analyzed.
- Chapter 6 presents a circuit conception and practical verification of the proposed methods. The concept of delayed feedback is verified using a discrete transimpedance amplifier. The other proposed methods are adapted to integrated circuit design and verified through post

layout simulation.

- Chapter 7 presents the final discussion and conclusions. Some guidelines for future work are also presented.

2 GAIN-BANDWIDTH OPTIMIZATION TECHNIQUES

“We were in a large room. Full of people. All kinds. And they had arrived at the same time. And they were all free and they were all asking themselves the same question: What is behind that curtain? They were all free. And they were all wondering what would happen next.” - Laurie Anderson.

Summary

This chapter addresses the problem of GBW maximization in amplifying circuits. GBW maximization is a old and well covered topic in circuit design. The foundations for GBW design come from analytic conditions originally reported by Bode. Techniques to reach or even overcome these limitations were proposed throughout the years. These techniques can be divided into two fundamental categories: conceptual – when the improvements result from a general concept applied to an amplifier; and circuit level – when the improvements result from the usage of special transistorized amplifier arrangements. This chapter discusses techniques belonging to the two above categories, exploring their most relevant features and drawbacks.

2.1 Preliminary Considerations

The main objective of any amplifier circuit is to guarantee as ideally as possible a targeted input-output relation. There are four different kinds of amplifiers: voltage amplifiers – when the input and output signals are both voltages; current amplifiers – when both input and output signals are currents; transimpedance amplifiers – when the input and output signals are respectively a current and a voltage; and transadmittance amplifiers - when the input and output signals are respectively a voltage and a current. The relation between the input and the output signals is the gain of the amplifier, defined as the ratio of the output variable to the input variable. The gain of the amplifier is generally one of the required design parameters.

Another important requirement is the amplifier bandwidth, that is, the amplifier's ability to maintain the required gain over a specified frequency range. Both gain and bandwidth are generally limited by the elements that compose the amplifier circuit. In general amplifiers must rely on transistors as active elements in order to provide a mean to achieve the required gain, so ultimately, transistors and their interconnections are the principal source of gain and bandwidth limitations (Raman amplifiers are exceptions to this description, since in both cases the amplification mechanism does not rely on transistors or any other active element [266]).

One important measure of the gain and bandwidth capabilities of an amplifier is its gain-bandwidth product or GBW for short. Maximization of the GBW is an amply debated problem in amplifier design. The question of how large can the bandwidth be, for some prescribed gain requirement on a given amplifier circuit has been amply discussed since the early ages of electronic design. Bode was the first to formalize the problem and to present a solution [248]. Bode's approach reveal that the inter-terminal capacitances of the active elements composing the amplifier circuit are the main causes of GBW limitation. This result was already known amongst filter designers, like Wheeler [23]. However, Bode was the first to provide a theoretical proof based on the analytic properties of realizable transfer functions. According to Bode's result the maximum GBW of an amplifier is fixed by the inter-terminal capacitances of the active devices. The significance of this result had a paramount impact on circuit design. Disclosing the limits provided insight on the extreme GBW capabilities and also ways to approach this limit. A whole set of new strategies was proposed to approach and even surpass Bode's limits. Distributed and chain amplification figure amongst the most relevant design strategies with these aims.

GBW maximization techniques have evolved trough the years following different design perspectives. It is possible to maximize the GBW of one amplifier taking three different approaches: increasing the gain, increasing the bandwidth or, more generally, increasing both gain

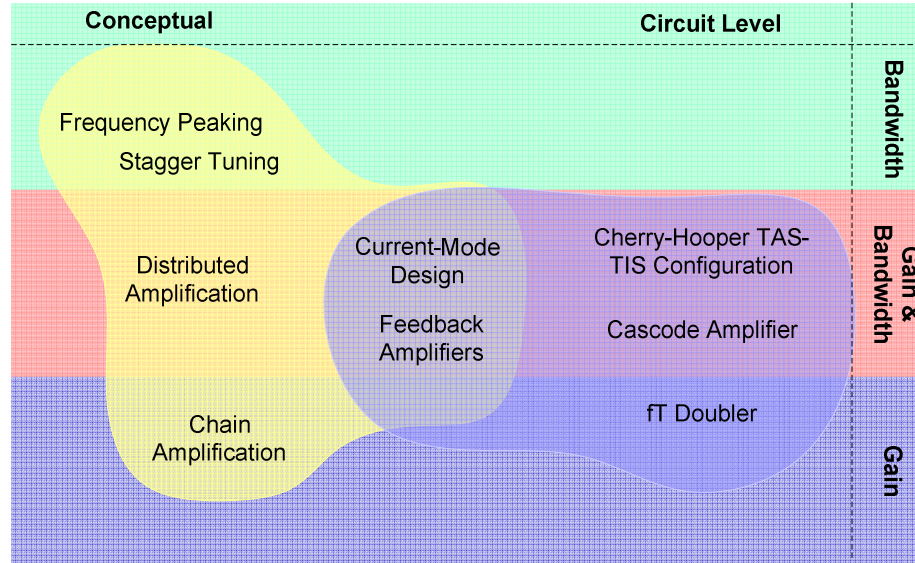


Figure 2.1 - GBW optimization techniques.

and bandwidth. These different approaches are implemented using several known circuit design techniques. These techniques can be further classified into two general categories:

- 1) Conceptual – when dealing with general design concepts that are not restricted to some special class of circuits;
- 2) Circuit-level – reporting to some special circuit arrangements known to have superior GBW characteristics.

Figure 2.1 shows several GBW optimization techniques organized according to the above criteria; optimization strategy (vertically stacked) and complexity level (using Venn diagrams). According to Figure 2.1 conceptual techniques covers the entire range of optimization strategies, with: i) frequency peaking and stagger tuning as bandwidth optimizing techniques; ii) distributed amplification, current-mode design and feedback amplifiers as gain and bandwidth optimizing techniques; and iii) chain amplification as gain optimizing techniques. Current-mode design and feedback amplifiers are mainly conceptual design techniques, nevertheless their implementation can be very dependent on circuit design considerations, thus, these two techniques are also classified as circuit level techniques. Finally, the Cherry-Hooper TAS-TIS configuration, the cascode amplifier and the f_T doubler are clearly circuit level design techniques, adequate to both gain and bandwidth optimization, except for the f_T doubler which is mainly a gain optimization technique.

The following sections explore the details of the general GBW limitations on amplifying circuits, and present an overview of the aforementioned GBW optimizing techniques. Due to the general character and importance of conceptual techniques, these are discussed in separated sections. Circuit-level techniques (except for feedback amplifiers) are discussed in a single section.

2.2 General Gain-Bandwidth Limitations

General gain-bandwidth product (GBW) limitations rise from common characteristics on the realizability of stable network functions. A stable network function, which can be some transfer function of a given type or some driving point impedance, has the unique property of being analytic inside the complex right half plane (RHP). This means that all the singularity points are located on the complex left half plane (LHP). As a general consequence from the Cauchy-Goursat theorem [248, 286], the contour integral of any such function for any arbitrary path inside the complex RHP must evaluate to zero, or equivalently,

$$\oint_C F(s) ds = 0 \quad (2.1)$$

where $F(s)$ is analytic inside the closed contour C . Using this result, Bode developed a wide variety of restrictions on network functions, using some special forms of integrand functions [248]. For the GBW analysis one such form is of particular interest.

Considering a complex function of $s = \sigma + j\omega$, $F(s) = X(s) + jY(s)$, satisfying the following restrictions:

1. $X(\omega)$ is an even function of ω .
2. $Y(\omega)$ is an odd function of ω .
3. $F(s)$ has no singularities in the interior of the RHP.
4. $F(s)$ can have logarithmic singularities and branch points³ on the $j\omega$ axis, but no poles.
5. $F(s)/s$ vanishes when s is made arbitrarily large.
6. $F(s)$ approaches $X_\infty + jY_\infty/\omega$ as ω approaches infinity, respectively. X_∞ and Y_∞ are real quantities.

Examples of such functions are passive impedances of minimum reactance type or transfer functions of minimum phase shift type⁴.

Let Figure 2.2 represent the integration path for the evaluation of (2.1). Any function vanishing more rapidly than $1/\omega$, will of course, lead to a zero contribution around the semicircular part of the contour, as the radius is made arbitrarily large. This implies that the total contour integral can be evaluated considering only the $j\omega$ axis contribution, and that the final result should be equal to zero.

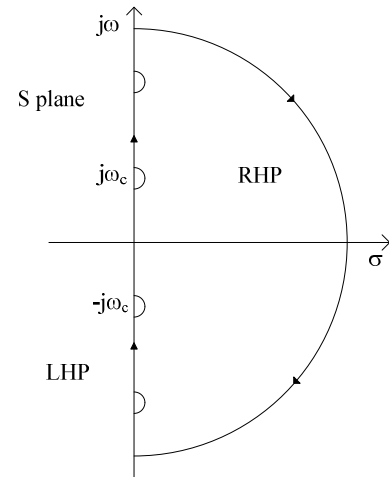


Figure 2.2 - Integration contour.

³ Branch points are singular points where the real and imaginary parts of $F(s)$ interchange with each other.

⁴ Driving point impedances or transfer functions having only poles and zeros with negative real part.

From the GBW perspective it is important to find the restrictions on gain for a fixed bandwidth. For this purpose Bode considered one particular form satisfying all the above restrictions,

$$\oint_C \frac{F(s) - X_\infty}{\sqrt{1 - \omega^2/\omega_o^2}} ds = 0 \quad (2.2)$$

Where it is assumed that $(1 - \omega^2/\omega_o^2)^{1/2}$ is a positive real quantity for $-\omega_o < \omega < \omega_o$, a positive imaginary for $\omega > \omega_o$ and a negative imaginary for $\omega < -\omega_o$. Since the semicircular portion of the path vanishes as the radius is made arbitrarily large, the integral from 0 to infinity of the even part of the integrand must also vanish. However, since $(X(\omega) - X_\infty)/(1 - \omega^2/\omega_o^2)^{1/2}$ is even for $|\omega| < \omega_o$ and odd thereafter, and $jY(\omega)/(1 - \omega^2/\omega_o^2)^{1/2}$ is even for $|\omega| > \omega_o$ and odd for smaller values of ω , it must result that,

$$\int_0^{\omega_o} \frac{X(\omega) - X_\infty}{\sqrt{1 - \omega^2/\omega_o^2}} d\omega = - \int_{\omega_o}^{\infty} \frac{Y(\omega)}{\sqrt{\omega^2/\omega_o^2 - 1}} d\omega \quad (2.3)$$

This result specifies the restrictions of the real part of $F(s)$ in a finite bandwidth if its imaginary component is to be prescribed outside this frequency range.

Bode applied successfully equation (2.3) to circuit design in various ways. It is outside the scope of this work to present these design strategies. However, the usage of (2.3) allows to disclose important GBW restrictions on circuit design. The next sections explore these limitations, for both two-terminal and four-terminal interstages.

2.2.1 Two-Terminal Interstages

Two-terminal interstage structures are perhaps the most widely applied topologies in the design of multistage amplifiers. Figure 2.3 depicts the two-terminal interstage topology. The connection

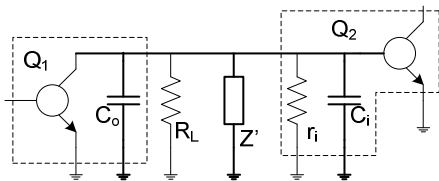


Figure 2.3 - Two-terminal interstage design⁵.

between two neighbor active devices is made directly. This connection may include several parallel elements to consider: the parasitic capacitances, C_o and C_i (due to active devices and line extension), some parallel resistance, R_L (mainly load resistance used to establish the stage's gain - input and output resistances from the active devices, act as reducing elements on the total load of the stage), and a matching impedance, Z' . The matching impedance can be used to further extend the bandwidth capabilities of the amplifier. Inductance shunt peaking techniques [51, 55, 56, 60] are an immediate example of this matching scheme. In order to simplify the formulation, it may be assumed that the total parallel impedance, $Z(j\omega)$, includes both r_i , R_L and $Z'(j\omega)$. Also, the two

⁵ Figure 2.3 uses a generic transistor definition. For further information on this see appendix A 1.

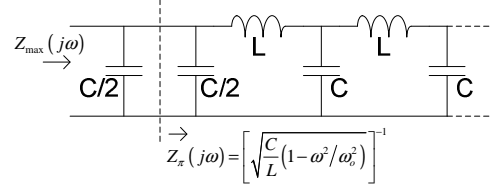


Figure 2.4 - Constant k low-pass filter.

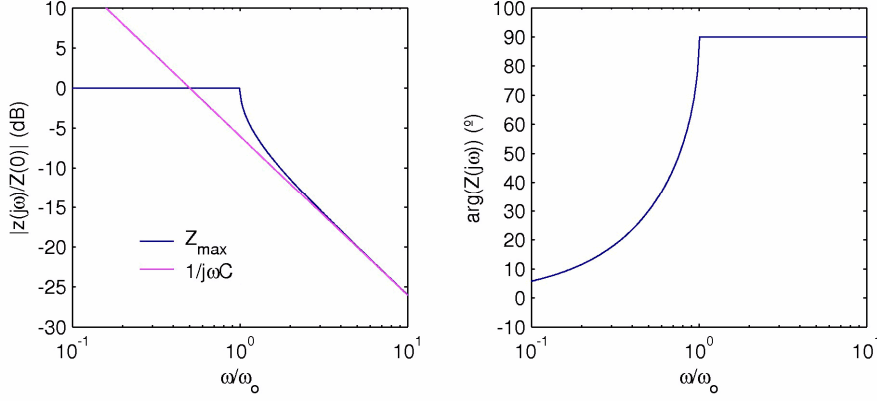


Figure 2.5 - Impedance giving the maximum flat gain: amplitude and phase.

capacitances C_o and C_i in Figure 2.3, add together to form a single capacitance C .

From the amplifier design perspective it is important to know what would be the maximum achievable flat gain over a restricted bandwidth. In other words: what would be the maximum GBW for an interstage having flat frequency response over a limited frequency range. The solution to this problem was empirically known prior to Bode's results [248]. Filter designers had applied similar results to produce filters with nearly optimum frequency response characteristics. According to these results the largest possible resistance that can be maintained over a prescribed bandwidth corresponds to the driving point impedance of an ideally selective constant k low-pass filter terminated with a full-shunt at the driving end, as depicted in Figure 2.4 and given as,

$$Z_{max}(j\omega) = \left[\frac{C}{2} \left(\sqrt{\omega_o^2 - \omega^2} + j\omega \right) \right]^{-1} \quad (2.4)$$

Figure 2.5 shows the amplitude and phase response of a normalized $Z_{max}(j\omega)$, assuming a unitary cut-off frequency. At the cut-off the height of $Z_{max}(j\omega)$ is 6dB above the level due to one capacitance alone. The high frequency behavior of $Z_{max}(j\omega)$ degenerates into the impedance of one capacitance.

Bode compared $Z_{max}(j\omega)$ against the general interstage impedance $Z(j\omega)$ using equation (2.3) to show that, this was the interstage impedance which would provide the maximum GBW. Knowing that g_m represents the transconductance of Q_1 and $Z(0)$ represents the magnitude of $Z(j\omega)$ at the origin, the gain of the interstage of Figure 2.3 is simply $g_m Z(0)$. Using the cut-off frequency ω_o to

represent the bandwidth of the interstage, Bode showed that,

$$g_m Z(0) \omega_0 \leq 2 \frac{g_m}{C} \quad (2.5)$$

The left side of (2.5) represents the GBW of the interstage, the factor $2g_m/C$ is the general restriction on GBW for two-terminal interstages. Considering the simple RC coupled interstage (2.5) represents a GBW improvement by a factor of 2, which shows that the maximum bandwidth improvement, for the same level of gain and same capacitances C_i and C_o , is also 2.

This result was previously reported by Wheeler in 1939 [23]. Its theoretical demonstration was issued only in 1945 due to Bode [248] and independently by Hansen [24]. Hansen [24] did go one step further, considering that the interstage can be of band-pass type and that GBW is restricted by the same factor, $2g_m/C$.

The factor $2g_m/C$ can be easily related with the transition frequency of an active device. If C_i and C_o represent the transistor's equivalent input and output capacitance respectively, using the intrinsic capacitances c_x (between terminals Y and X, see appendix A 1 for more details) and c_z (between terminals Y and Z) and the voltage gain factor given as $g_m Z(0)$, results in,

$$\begin{aligned} C_i &= c_x + c_z (1 + g_m Z(0)) \\ C_o &= c_z \left(1 + \frac{1}{g_m Z(0)} \right) \end{aligned} \quad (2.6)$$

Obviously, both C_i and C_o are affected by Miller effect, due to the fact that c_z forms a feedback path between input and output. Using (2.6) in (2.5) results in,

$$g_m Z(0) \omega_0 \leq \frac{2\omega_T}{M(1 + g_m Z(0) + 1/g_m Z(0))} \quad (2.7)$$

where ω_T is the transition frequency in rad/s given by $g_m/(c_x + c_z)$ and $M(u)$ is a real valued function quantifying the Miller effect, define by,

$$M(u) = 1 + uc_z / (c_z + c_x) \quad (2.8)$$

Equation (2.7) shows that the maximum GBW that can be realized with a two-terminal interstage is essentially restricted by $2\omega_T$, since for real transistors c_x is usually larger than c_z by a factor larger than the voltage gain of the stage.

2.2.2 Four-Terminal Interstages

Four-terminal interstages are attractive to amplifier design due to their ability of providing even more gain than two-terminal interstages. The concept behind four-terminal interstage design is depicted on Figure 2.6. The connection between two stages is accomplished using a four-terminal

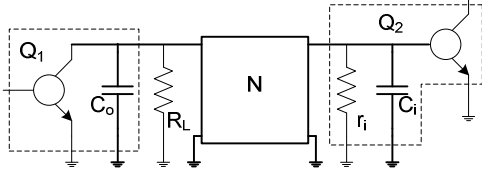


Figure 2.6 - Four-terminal interstage design.

network N . This connection also includes: two parasitic capacitances, due to the input (C_i) and output (C_o) of the two active devices; and two resistances, due to stage gain assignment (R_L) and input of the active device (r_i). For the four-terminal interstages case there

is no restriction on the maximum phase variation as for two-terminal interstages (restricted to 90°). This allows more design freedom, and also the advantage of larger GBW values. The nature of GBW restrictions are intimately related with the inter-terminal parasitic capacitances [27, 248]. For a four-terminal interstage having a common ground terminal, there are three such restrictions. Following the results of Bode [248], Hansen [23] and Spilker [27], the GBW limitations of the four-terminal interstage of Figure 2.6, having low-pass constant gain response are stated as

$$g_m Z_T(0) \omega_o \leq \min \left\{ \frac{\pi^2}{4} \frac{g_m}{C_i}, \frac{\pi^2}{4} \frac{g_m}{C_o} \right\} \quad (2.9)$$

Two advantages of four-terminal interstages are evident; the first is that four-terminal interstages can provide higher GBW than two-terminal structures – the GBW limitation is for this case 2.47 times the ratio between the active device's g_m and its larger parasitic capacitance (usually C_i). The second advantage is their ability to effectively decouple the parasitic capacitances of the two active devices interconnected, which is also an explanation for larger GBW. Bode also showed that the theoretical limit can be effectively approached if C_i and C_o have equal values [248]. This interesting result stems directly from (2.9) multiplying and dividing simultaneously by $C_i + C_o$,

$$g_m Z_T(0) \omega_o \leq \frac{\pi^2}{4} \frac{g_m}{C_i + C_o} \left(1 + \frac{C_o}{C_i} \right) \quad (2.10)$$

According to (2.5) taking C_i and C_o equal to $C/2$, the maximum GBW that can be obtained from a four-terminal interstage is 2.47 higher than that which would be obtained with a two-terminal structure having the same prescribed parasitic capacitance. This also means that even for situations where C_i and C_o are very different from each other (like in real transistorized circuits, due to Miller effect and asymmetries within the transistor model), it is possible to improve the GBW adding some external capacitance to C_o (the smallest capacitance of the interstage).

An interesting result can be derived if (2.10) is compared to the transition frequency. Starting from (2.9) and taking C_i as the most restrictive capacitance, the final GBW is,

$$g_m Z_T(0) \omega_o \leq \frac{\pi^2}{4} \frac{\omega_T}{M(g_m Z(0))} \quad (2.11)$$

For the band-pass type, a four-terminal interstage can provide slightly more gain than for the

low-pass type. Hansen [24] proved that for this case the GBW limitation can reach,

$$g_m Z_T(0) \omega_o \leq 2.53 \frac{\omega_r}{M(g_m Z(0))} \quad (2.12)$$

which is slightly more than the $2.47 (\pi^2/4)$ for the low-pass case.

The work developed by Bode [248], Hansen [24] and Spilker [27] presented and fully explained the GBW limitations on interstage networks. Nevertheless, several cases were left uncovered. For instance, the possibility to have any type of terminating impedance rather than a parasitic capacitance, or even interstages including active elements. The first case was attacked by Fano [25, 246], using a generalization of Bode's results applied to the reflection coefficient due to arbitrary impedance terminations. Several years later, in 1964, Youla issued a generalization able to deal with second problem [29]. Repercussions of Bode-Fano-Youla results find applications in many presently debated problems, irrespectively to the many technological improvements that became accessible all over the past years; in particular, the problem of GBW maximization was and still is a fundamental problem on amplifier design. Several extensions and modifications of the original Bode results were meanwhile advanced [26-28].

2.3 Chain Amplifiers

The problem of designing amplifiers having high gain over a large bandwidth is not new. As the results of the previous section showed, there are always restrictions on the achievable GBW of the amplifier. These restrictions arise due to the presence of "parasitic impedances" in the input and output ports of amplifiers. These parasitic impedances are mainly due to the intrinsic capacitances of the active devices and transducers. In past years, these parasitic capacitances had large values due to the large dimensions of the thermionic valves. Nowadays, the availability of faster and smaller transistors has not reduced this problem. The constant seek of larger and larger system bandwidths, which justify the constant demanding of amplifiers able to deal with small signal powers over larger bandwidths, has maintained the relative challenge. Most of the presently designed amplifiers are integrated amplifiers. The interconnection medium inside an integrated circuit (IC) is predominantly capacitive; meaning that parasitic capacitances exist throughout the IC, originating not only by transistors.

An efficient way of achieving high gain in amplifiers is to use cascade connections of several amplifying stages. Under the assumption that the frequency response of each stage should be flat inside some prescribed bandwidth, the overall gain is set by the product of each stage's gain. This assumption simplifies the evaluation of the bandwidth of the total cascade. However, this is not the unique possibility to achieve an overall flat gain for a cascade association of several stages: for

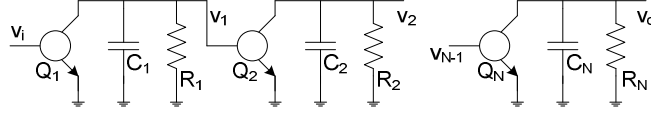


Figure 2.7 - RC coupled chain amplifier.

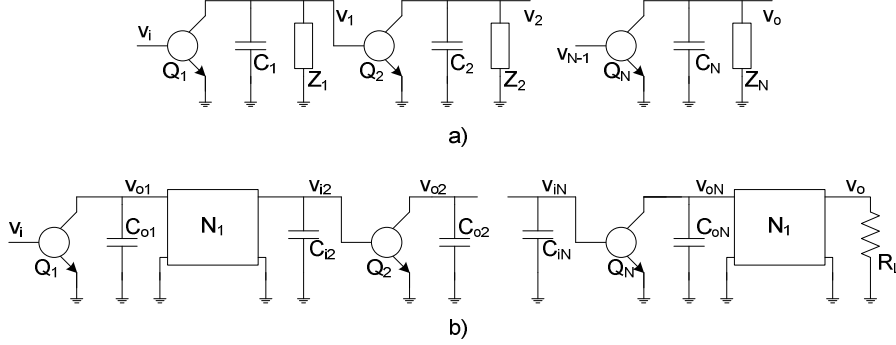


Figure 2.8 - Chain amplifiers: a) with two-terminal interstages; b) with four-terminal interstages.

instance, it is possible to extend the bandwidth of a low-pass amplifier, using one or more cascaded tuned amplifiers having pass-band with high gain near the cut-off of the first stage. This is known as stagger-tuning, after Wallman [34, 35], and recent examples of this technique can be found in [36, 37, 41].

In a cascade amplifier the output of one stage attacks the input of the next stage in the chain, and for this reason these amplifiers are known as chain amplifiers. One important limitation of chain amplifiers is the fact that increasing the number of stages in the chain improves the gain but also degrades the bandwidth. This phenomenon can be best understood considering the simple case of an RC coupled chain amplifier (Figure 2.7). Assuming that the stages in the chain have equal gain and bandwidth, the overall bandwidth is given by,

$$\omega_c = \omega_s \sqrt{\sqrt{N} - 1} \quad (2.13)$$

This well known result states that for a chain amplifier composed of N identical stages, the cut-off frequency of the chain decreases with the number of stages. When the stages have different cut-off frequencies this effect still persists. A fast rule of the thumb for these cases is to use the geometric mean of the cut-off frequencies.

Considering that the interstages used can be of any type rather than simple RC coupled circuits, the GBW potentialities of the chain amplifier can be enhanced. There are two possibilities to design chain amplifiers using general interstage design as discussed, the first uses two-terminal interstage structures between neighbor stages; the second uses four-terminal interstage structures instead. The GBW possibilities of second case are obviously less restricted than the first. Figure 2.8 depicts the two mentioned approaches. The two-terminal interstage topology represented in Figure 2.8a, has each interstage composed by a single capacitance C_k and one matching impedance Z_k ; obviously, C_k represents the sum of all shunting capacitances attached to the interstage node. Z_k may include both

the loading resistance and the input resistance of the stage.

Starting with the two-terminal interstage chain amplifier, Linvill [30] showed that the maximum flat gain that can be achieved over a constant bandwidth ω_b with a chain amplifier using this kind of interstage is given as,

$$G_o = \prod_{k=1}^N g_{m_k} Z_k(0) \leq \prod_{k=1}^N \left(2 \frac{g_{m_k}}{\omega_o C_k} \right) \quad (2.14)$$

This is simply the product of the maximum flat gain restrictions of each stage assuming a fixed bandwidth of ω_b . Assuming that each stage is designed so as to attain the limit case (constant gain inside the prescribed bandwidth ω_b), the loading between stages does not affect the bandwidth of the chain and the overall bandwidth is exactly ω_o ⁶. Assuming also that each stage is designed so as to match the same gain per stage as the reference RC couple chain amplifier, it follows that the same gain restrictions must hold for each stage in both cases. Since for the simple RC coupled case each stage has a bandwidth $\omega_k = 1/R_k C_k$ the gain restrictions are,

$$g_{m_k} R_k = \frac{g_{m_k}}{\omega_s C_k} \quad (2.15)$$

Comparing (2.15) with the corresponding gain restriction of the two-terminal chain amplifier (each product term in (2.14)), follows that the bandwidth relation for each stage is $\omega_b = 2\omega_s$. Under these considerations the GBW of the chain amplifier using two-terminal interstages is simply given by $2G_o\omega_s$. Since the overall gain of the two topologies is assumed equal, the resulting bandwidth enhancement ratio (BWER) is,

$$BWER_{II} \leq \frac{2}{\sqrt[N]{2} - 1} \quad (2.16)$$

For the four-terminal case (Figure 2.8b), there are two different strategies to combat GBW limitations. One possibility is to design all the stages in the chain with optimally flat frequency response; the total chain is limited by the Bode's GBW limitation [248]. On the other hand, using staggering (in its more general form, stagger dumped tuning [35]), the individual stages may exhibit band-pass characteristics, and the GBW limitation is given by Hansen's result [24].

Assuming for simplicity sake that all stages are designed using simple low-pass four-terminal interstages with equal gain and bandwidth characteristics and following similar considerations as in

⁶ This is not the unique possibility to attain maximum flat response of the overall chain. Recalling that the very same limitation for two-terminal interstages also applies for band-pass structures [24], the chain amplifier can be designed more freely if all the interstages act in conjunction to produce an overall flat frequency response (as in stagger-tuned amplifiers [34]).

the above two-terminal chain amplifier, the resulting BWER is given by,

$$BWER_{IV} \leq \frac{2.47}{\sqrt{\sqrt[n]{2}-1}} \left(1 + \frac{C_{ok}}{C_{ik}} \right) \quad (2.17)$$

The ratio of C_{ok} to C_{ik} in (2.17) is equal for all the stages in the chain due to the equal gain and bandwidth assumption. It is also implicit that $C_{ok}+C_{ik}$ matched the interstage capacitance, C_k of the RC coupled chain for comparison purposes.

Equations (2.16) and (2.17) show clearly that there is much to gain in bandwidth by appropriate design of the interstages in a chain amplifier. These results are also valid for general single stage

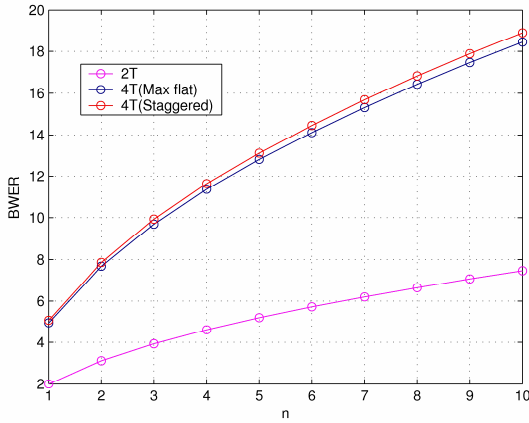


Figure 2.9 - BWER functions for two-terminal (2T) and four-terminal (4T) interstages.

amplifiers and will be applied in a section 2.6, concerning frequency peaking techniques. Figure 2.9 depicts the BWER functions for three different interstages configurations: i) using simple two-terminal (2T) interstages; ii) using four-terminal interstages (4T) with maximum flat gain; and iii) using staggered four-terminal interstages (the correspondent BWER is given by (2.17) changing the 2.47 factor by 2.53). It was also assumed for the 4T structures that $C_i=C_o$ and all the stages follow the same design strategy (maximum flat gain or

staggered). The benefit in BWER between 2T and 4T interstages is evident. As the number of stages increase the BWER increase rapidly for all the considered possibilities. Choosing between 4T interstages of maximum flat gain or staggered type may represent a slight benefit in BWER. Usually the two approaches are unavoidably employed in the same chain, and the real maximum BWER lies between the two.

On the above considerations it was assumed that the input and output circuits of the amplifiers did not pose severe restrictions on bandwidth. In general, the presence of an input capacitance does not affect the GBW when a signal voltage is applied directly to it. However, considering that the signal source is not ideal, the presence of internal impedances may affect the overall GBW. A more serious situation exists when the signal source is a current source. In this case the situation is identical to a two-terminal interstage. This is most relevant when transimpedance amplifiers are used to convert a signal current into an output voltage. In particular, if the current source is a photodiode with high intrinsic capacitance, this, together with the input resistance of the amplifier results in a severe input bandwidth restriction. These input limitations can be conveniently treated

using the same techniques employed in two-terminal and four-terminal interstages; in fact the input circuitry may be regarded as another matching interstage.

The output circuit of a chain amplifier also poses some restrictions on GBW, however this can be simply introduced in the above equations. For the two-terminal interstage case, this is already part of the result. For the four-terminal case, the last stage must be taken with special care. In Figure 2.8b, since the output of the last interstage is connected directly to a load resistance and no capacitance exists on the output node, there is no restriction posed by R_L . If however, some capacitance exists on this node (let it be C_L), its integral resistance limitation should be compared against the integral resistance limitation of the input port of the same interstage (as in equation (2.9)). Finally, the most restrictive should be taken. In deriving (2.17) it was assumed that the interstages have identical input and output capacitances, being the output capacitance of each interstage the largest one.

The impact of these results is evidenced by the published works of Hajimiri and Analui [31, 32], Lu *et al* [33], and Lee *et al* [59] both in 2004. Their results point towards the bandwidth advantages of using multi-pole networks between stages of amplification. This multi-pole technique is clearly covered by the more general four-terminal interstage chain amplifier design technique. The limitations claimed by Hajimiri and Analui agree with the above mentioned results.

2.4 Stagger-Tuned Amplifiers

Stagger tuned amplifiers were invented in 1948 by Wallman [34, 35]. The staggering technique consists in broadening the band of a tuned multistage amplifier by an appropriate de-tuning of its stages. Figure 2.10 depicts two typical interstage networks used in tuned amplifiers. The single

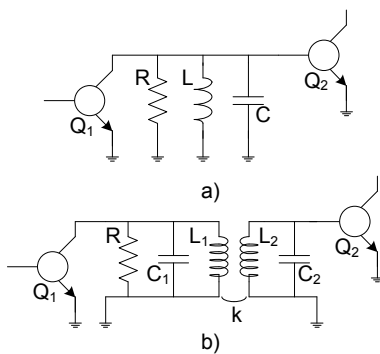


Figure 2.10 - Stagger tuned amplifiers: a) single tuned; b) double tuned.

tuned circuit uses only one inductance in parallel with load and parasitic capacitance of the stage. Double tuned circuit employs transformers; the tuning action is dependent on both sides of the transformer. Double tuned circuits are currently less interesting than their single tuned versions because of the transformer. It is not usual to find transformers with good quality associated to modern integrated circuit technologies, albeit some attempts to design and use transformers inside integrated

circuits have been reported [38-40, 57-59]. On the other hand, reasonable quality inductances are becoming mainstream components in high frequency IC (Integrated Circuit) technologies, making their use in electronic circuits less prohibitive than transformers.

Stagger tuned amplifiers can be used in cascade with low-pass amplifiers in order to enhance the

bandwidth of the overall amplifier. Examples of this strategy were reported by Honjo in [36, 37] and Shekhar in [41, 63]. Considering the circuit of Figure 2.10a, where R represents the load resistance used to set the gain and C represents the total parasitic impedance of the stage, the frequency response between the input of the first transistor to its output is given by,

$$H(j\omega) = \frac{g_m R}{1 + jQ \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)} \quad (2.18)$$

Where $Q = \omega_o RC$ is the quality factor, and $\omega_o = (LC)^{-1/2}$ is the resonance frequency of the RLC interstage. Using a frequency transformation from band-pass to low-pass (2.18) becomes,

$$H(ju) = \frac{g_m R}{1 + ju} \quad (2.19)$$

Now assuming a chain of identical stages with frequency response specified by (2.18), given the correspondence with (2.19), the bandwidth decreases with the number of stages in the same fashion as in (2.13). This means the frequency response will become more selective as the number of stages is increased.

One possible way to break this behavior is to de-tune the resonant frequencies of the individual stages. If this is done using Butterworth polynomials (see appendix A 2 as reference), the final result is a flat band-pass frequency response. Returning to the low-pass equivalent of (2.19), a product of N such transfer functions produce an overall transfer function with a N^{th} order denominator polynomial in u frequency variable. Choosing adequately the coefficients of this polynomial it is possible to attain a maximal flat response⁷. In order to apply this maximal flat procedure to (2.18) some considerations on its nature should be made, namely: (2.18) exhibits geometric symmetry - two frequencies ω and ω_o/ω produce the same amplitude; the Q factor controls the aperture of the characteristic; and the maximum value is attained at frequency ω_o and has the value $g_m R$.

It is reasonable to think that if the overall response is to be flat, at least the shape of the individual transfer functions should be maintained (it is possible to assume different shapes, however the tuning scheme is not so straight forward), thus the only variable parameters are the center frequency and gain of each stage; the center frequency is used to extend the band and the

⁷ By definition, a transfer function has maximal flat frequency response if its frequency response magnitude is a real and strictly decreasing in maximum sense, function of the frequency variable. For functions having polynomial forms, the maximal flat condition implies that these are also Butterworth polynomials. Further reference to this subject can be found in appendix A 2.

gain to compensate the loss produced on other stages.

Taking the case of two stages, maintaining the shape of the characteristic corresponds to maintain the Q value for all stages, furthermore, assuming that the center frequencies of the two stages act in such a way to produce a unitary geometric mean, that is: one stage is resonant at α and the other at $1/\alpha$. The composed transfer function can be written as,

$$H_2(j\omega) = \left(\frac{g_{m_1}}{C_1}\right) \left(\frac{g_{m_2}}{C_2}\right) \frac{1/\alpha}{d + j\left(\frac{\omega}{\alpha} - \frac{\alpha}{\omega}\right)} \frac{\alpha}{d + j\left(\omega\alpha - \frac{1}{\omega\alpha}\right)} \quad (2.20)$$

Where $d=1/Q$, and the terms g_{mk}/C_k are due to the transistor transconductance and the stage capacitance. Multiplying the denominator factors,

$$d^2 + \alpha^2 + \frac{1}{\alpha^2} + jd\left(\alpha + \frac{1}{\alpha}\right)\left(\omega - \frac{1}{\omega}\right) - \left(\omega^2 + \frac{1}{\omega^2}\right) \quad (2.21)$$

Observing in (2.21) that, $x^2 + x^{-2} = (x - x^{-1})^2 + 2$,

$$H_2(j\omega) = \frac{\left(\frac{g_{m_1}}{C_1}\right) \left(\frac{g_{m_2}}{C_2}\right)}{d^2 + \left(\alpha - \frac{1}{\alpha}\right)^2 + jd\left(\alpha + \frac{1}{\alpha}\right)\left(\omega - \frac{1}{\omega}\right) - \left(\omega - \frac{1}{\omega}\right)^2} \quad (2.22)$$

If d and α are adjusted in order to produce a Butterworth denominator polynomial (see appendix A 2) the magnitude of (2.22) becomes maximally flat. This objective is accomplished if the denominator of $H_2(j\omega)$ is compared to $B_2(u)$ (second order Butterworth polynomial in u). Using the appropriate frequency transformation, this corresponds to,

$$B_2\left[j\left(\omega - \frac{1}{\omega}\right)\right] = \delta^2 + j2\delta \sin\left(\frac{\pi}{4}\right)\left(\omega - \frac{1}{\omega}\right) - \left(\omega - \frac{1}{\omega}\right)^2 \quad (2.23)$$

Where $\delta = \Delta\omega/\omega_c$ the ratio between the bandwidth of the overall amplifier to its center frequency. Comparing the coefficients of the denominator of (2.22) with (2.23), results

$$\begin{cases} d^2 = \frac{\delta^2 + 4}{2} - \frac{1}{2}\sqrt{\delta^4 + 16} \\ \left(\alpha - \frac{1}{\alpha}\right)^2 + d^2 = \delta^2 \end{cases} \quad (2.24)$$

The values of the elements for each stage are easily related to parameters d and α using (2.18) and the accompanying discussion. This procedure can be generalized for an arbitrary number of stages, resulting in a set of equations like (2.24) for each two stages. If the number of stages is odd, then according to the brief discussion on Butterworth polynomials presented on appendix A 2 there will be a first order term with an independent set of equations. The general expressions for an

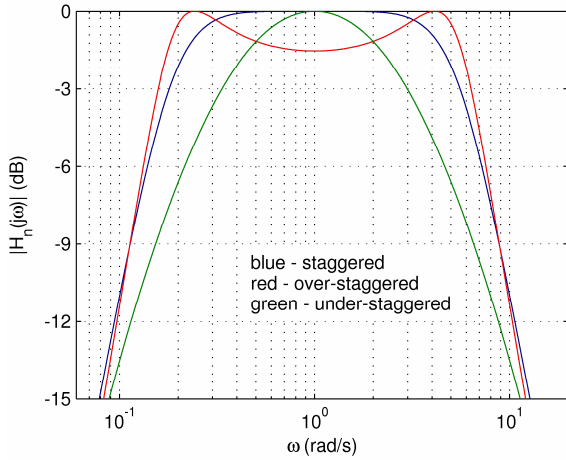


Figure 2.11 - 2 stages staggered amplifier frequency response.

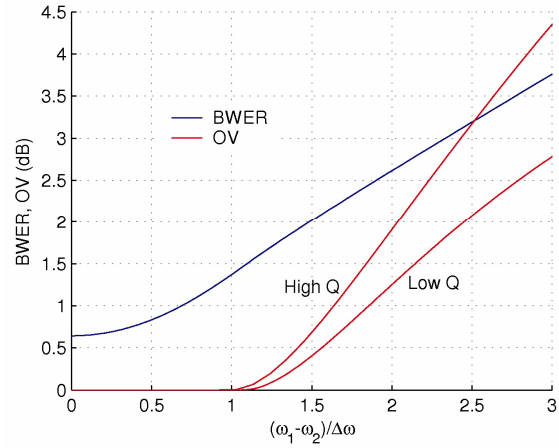


Figure 2.12 - 2 stages staggered amplifier, BWER and overshoot (OV).

arbitrary number of stages can be found in [34].

Larger bandwidths can be attained using stagger tuning if some gain variation can be tolerated (see Figure 2.11 and Figure 2.12). This is easily understandable, as staggering consists in separating the resonant frequencies of cascaded tuned amplifiers. This procedure can produce different frequency responses. If the circuit elements are adjusted as above mentioned, the overall characteristic is maximally flat. However, if the elements have such values that the maximally flat condition is not reached (the resonant frequencies are too close together), the frequency response resembles the frequency response of a tuned amplifier with broader bandwidth. Another possibility is to surpass the maximally flat condition (the resonant frequencies become seriously separated), for this case, the frequency response exhibits multiple maxima inside the useful bandwidth. These three situations are known as staggered, under-staggered and over-staggered, respectively. Figure 2.11 depicts these three situations for a case where the resonant frequencies remain fixed and Q is varied from its ideal staggered value. Over-staggering can be used to broadening the bandwidth of the amplifier. However this must be done with special care; if the maxima can attain magnitude larger than $3dB$ the amplifier exhibits multiple cut-offs. Figure 2.11 depicts this possibility for a two stage staggered amplifier. The staggering characteristic (blue line) shows the possible bandwidth enhancement against the normalized peak separation frequency. Both axes were normalized using the geometric mean bandwidth of the two stages. Bandwidth enhancement of more than 3 is possible for high Q amplifiers; low Q amplifiers can achieve even better results (as indicated by the overshoot – red lines). It is interesting to notice that the flat gain condition prevents reaching the maximum BWER of 7.86 given from (2.17) (with a factor 2.53) for a two stage chain amplifier.

The superiority of the stagger damped tuning technique over the simple stagger tuning scheme follows directly from a careful inspection of the circuit of Figure 2.10b. This circuit exhibit in a

single stage a double tuned circuit, thus the procedure of adjusting the resonant frequencies of two different stages just discussed can be applied to this case. Furthermore, the nature of the interstage also allows better GBW than simple stagger tuned stages. However due to the stringent requirement of having a transformer, the utility of this technique is of limited application in present integrated circuit design. Nevertheless, the general formulation for this scheme can be found in [35].

2.5 Distributed Amplifiers

The phenomenon of distributed amplification was first reported by Percival in 1935 [42]. In its first appearance distributed amplification consisted of a means of achieving higher frequencies in thermionic valves. Percival proposed an optimized way of designing the plate and grid arrangements in a valve, so that the plate and grid capacitances become uniformly distributed along the so called grid and plate lines. The grid and plate terminals were no longer single points; they became transmission lines, and the total grid and plate capacitances formed the shunt arms of those lines. By a careful adjustment of the distributed inductance and the capacitance of these lines, the properties of this new device could be adequately controlled and the cut-off frequency became simply determined by the cut-off frequencies of the lines (which were designed to have equal cut-offs).

Percival also reported that the same arrangement could be adapted for high bandwidth amplifier design, if the valves on the amplifier were arranged along two independent transmission lines. The first attempt to generalize and model the distributed amplification phenomenon was reported some years later in 1948 by Ginzton *et al* [43]. Ginzton thus baptized the phenomenon of distributed amplification and the distributed amplifier. Figure 2.13 shows the basic arrangement of a distributed amplifier. In a distributed amplifier several active devices (which may, or may not, be single transistors) are arranged in an additive fashion. The input terminals (Y terminals) form part of the input transmission line, composed of inductances L_i as series arms, and C_i as shunt arms (mainly due to the active device). One end of the input line is connected to the signal source and acts as

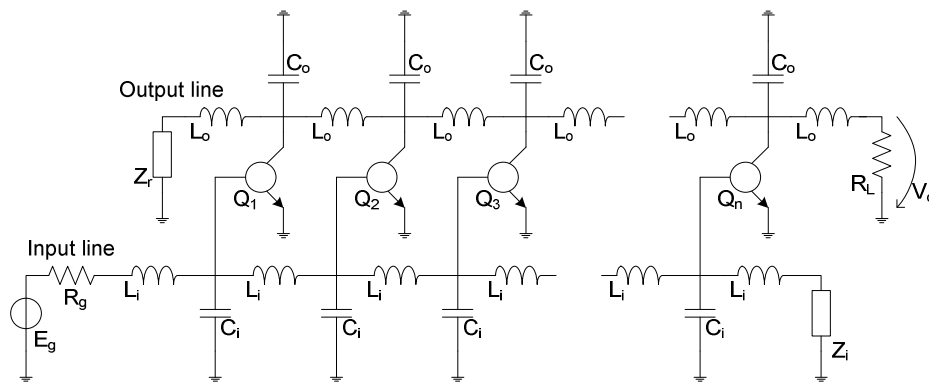


Figure 2.13 - Distributed amplifier.

input; the other end is adequately terminated in order to minimize reflections. The output terminals (Z terminals) form part of the output transmission line, composed of inductances L_o as series arms, and C_o as shunt arms. The end near the input circuitry is adequately terminated (the so called reverse termination). The other end acts as amplifier's output. The two lines are designed to exhibit equal propagation velocities.

A generator connected to the input port of the input line causes a signal wave to travel along the line. As the signal reaches the input of each device, a current is produced on the output terminal of the same device. This causes a signal wave to travel in both directions on the output line. The waves traveling towards the reverse termination are completely absorbed at the reverse load Z_r , thus preventing their interference to the output voltage. The waves traveling toward the load end of the line add together in phase. Thus the total output voltage is directly proportional to the number of active devices. No matter how low the gain of each active device maybe (even if it is less than unity), as long as the losses along the line are compensated, the signal on the output line will increase and can be made as large as desired, merely using an appropriate number of active devices. A simple inspection of the GBW potentialities of the distributed amplifier can be done assuming that [43]: i) there is no loss on both transmission lines; ii) the lines have equal cut-off frequencies, $\omega_c = 2(L_i C_i)^{-1/2} = 2(L_o C_o)^{-1/2}$, and hence the same propagation velocities; iii) it is possible to have matched terminations on the line and also that the line is perfectly matched between stages. Under these assumptions, a voltage E_g applied at the input line will cause a current $g_m E_g$ to flow on each active device. The output impedance presented to the output terminal of each active device is $Z_{o2}/2$ (where the subscript 2 indicates the output line), thus the output voltage due to one single device is $g_m Z_{o2} E_g / 2$. Since for the previous assumptions, the signals of each stage add together in phase at the output terminal, the voltage gain is given by,

$$A_v = \frac{V_o}{E_g} = \frac{n g_m Z_{o2}}{2} \quad (2.25)$$

using $Z_{o2} = (L_o / C_o)^{1/2}$ and the definition of ω_c ,

$$A_v = \frac{n g_m}{2} \sqrt{\frac{L_o}{C_o}} = \frac{n g_m}{2 C_o} \sqrt{L_o C_o} = n \frac{g_m}{\omega_c C_o} \quad (2.26)$$

where C_o represents the transistor's output capacitance. Assuming that C_o is approximated by the transistor's intrinsic capacitance c_z and introducing the transition frequency, ω_T in (2.26) gives,

$$A_v = n \frac{\omega_T}{\omega_c} \left(1 + \frac{c_x}{c_z} \right) \quad (2.27)$$

Equation (2.27) shows that the GBW of a distributed amplifier is indeed boundless; large GBW's

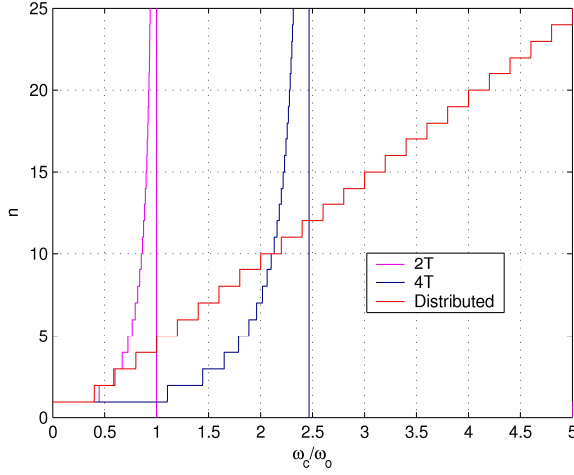


Figure 2.14 - Comparison of amplifying strategies.

can be achieved using a sufficiently large number of active devices. Figure 2.14 displays the comparison between the numbers of stages needed to fulfill an overall gain requirement, using chain amplifiers with two-terminal or four-terminal interstages and distributed amplifiers. The performance comparison assumes the same cut-off frequency ω_c , a constant gain condition (assuming a gain of 5) and $C_i=C_o=C/2$. Under this conditions the overall gain for all the three structures is directly dependent the common factor $\omega_o=2g_m/C$. Wheeler, named this frequency as the bandwidth index [23], representing the frequency condition where a two terminal interstage delivers unitary gain. The available gain per stage is a function of the ratio between ω_o and the stage cut-off frequency ω_c , as can easily be derived from the correspondent GBW restrictions. Obviously, ω_o is directly connected to the transition frequency of the active device (as shown previously). Figure 2.14 shows that the distributed amplifier is the unique solution to amplify signals with frequencies above $2.47\omega_o$. A careful inspection of the curves for two-terminal and four-terminal interstages based chain amplifiers, show that the number of stages increases sharply as the limit frequency is approached (ω_o and $2.47\omega_o$, respectively). This behavior is understandable, since approaching the limit frequency implies severe gain reductions, thus meaning that a larger number of stages will be needed to fulfill the overall gain requirement. Above the limit frequency these structures can not provide gain. The ideal distributed amplifier does not have this stringent limitation; the number of stages increases linearly with the cut-off frequency and gain above $2.47\omega_o$ can be provided. Furthermore, designing a chain amplifier with a large number of stages is always a difficult task. In common practice, this kind of structures can achieve at most 50% of the available frequency limit with a modest number of stages. On the other hand, distributed amplifiers are complicated to design, but their complexity does not increase much with increased number of stages.

The above assumptions on distributed amplifiers are far from common practice [44]. In practice the loss along the transmission line is of paramount importance, especially in current integrated circuits technologies. Integrated inductances have moderate to low Q , and thus the loss produced along the line has to be compensated with the gain of the active devices. On the other hand, the active devices may operate as “active” elements within a restricted frequency band: the maximum

oscillation frequency, f_{max} (see appendix A 1), is one that can not be surpassed. For frequencies above f_{max} , the active devices become “passive” elements, thus eliminating the loss compensation mechanism. These two factors, the losses along the line and f_{max} , provide a useful indicator for the maximum GBW that can be achieved with a distributed amplifier. Other limiting factors arise due to imperfect matching of the lines, and different propagation velocities. A detailed analysis of these effects can be found on [45-47, 266]. These factors contribute to destroy the linearized behavior depicted on Figure 2.14.

2.6 Frequency Peaking

Frequency peaking techniques are a means of extending the bandwidth of one amplifier, adding elements that can produce gain above the otherwise cut-off region. Usually this consists of adding resonant circuits, with resonance frequencies adequately tuned. Inductive peaking schemes [51, 53, 55-63], are perhaps the best choice to produce this desired frequency peaking. However, these are not the only means of achieving the desired peaking behavior. Some capacitive arrangements can also be used with the same goal, as shown in [64-74]. Strategies to shape the actual effect of the peaking circuit are also diverse; for instance, using maximally flat design procedures (recurring to Butterworth polynomials [286]) or maximally flat delay design (using Bessel polynomials [286]). The maximally flat Butterworth design strategy is adopted for the following discussion.

2.6.1 Butterworth Design

Butterworth polynomials are the best choice to achieve a maximum flat bandwidth operation. As presented on appendix A 2, Butterworth polynomials have a whole set of interesting characteristics, namely: i) they are strictly monotonic; ii) they have all the first $n-1$ derivatives identically zero at the origin, for a n^{th} order polynomial; iii) their roots lie on a circle of defined radius; and iv) they have constant cut-off frequency, set by the radius of the roots. These properties make Butterworth polynomials the elected choice for the synthesis of low-pass type transfer functions (using adequate frequency transformations it is also possible to synthesize other kinds of frequency responses, as for instance in stagger tuning). The strictly monotonic behavior is useful to the design of low-pass frequency responses having strictly decreasing character. The fact that for an n^{th} order polynomial, all its first $n-1$ derivatives are identically zero at the origin, adds also the unique property of being the maximal flat possible near the origin. Butterworth polynomials are the unique polynomials having this property: several other types of polynomials have the strictly monotonic character without this maximal flat property (see for instance Bessel polynomials [286], Papoulis polynomials [48, 49] and a general kind of polynomials discovered by Filanovsky [54], where all the strictly

monotonic polynomials can be included). Finally, the constant bandwidth property is of paramount importance for predicting system bandwidth; dealing with 1st and 2nd order systems is a simple task, but, as the number of poles in the system increases, predicting its cut-off characteristics can become a painful and time consuming task. Allocating the poles on a circle, according to a Butterworth design procedure simplifies this task, since the bandwidth is immediately defined [279, 284, 286].

It should be observed that larger bandwidths can be obtained if some gain variation inside the band can be tolerated. This is the very same argument observed in stagger tuned amplifiers; the over staggered possibility allows to further extend the amplifier bandwidth, until the 3dB maxima condition is observed. More conservative approaches limit the amount of gain variation to less than 3dB. This possibility of gain variation also has its price. An immediate consequence is that the phase response may become non-linear inside the useful band, thus impairing the usage of the amplifier. Non-linear phase responses imply delay frequency dependence; signals with different frequencies will be affected by what is called delay distortion [279, 284]. When maximal flat delay is a target goal, the design strategy should resort to Bessel polynomials, [286]. The bandwidth enhancement properties of this kind of polynomials are more modest.

Shaping the frequency response of one amplifier to match the maximal flat case usually relies on the right positioning of the system poles. Assuming that a general low-pass type transfer function has the form of,

$$H(s) = \frac{A_o}{D(s)} = \frac{A_o}{1 + a_1s + a_2s^2 + \dots + a_ns^n} \quad (2.28)$$

where A_o represents the zero frequency gain and a_k are the k^{th} order time constants of the system. Using the normalizing transformation for $D(s)$, $s = u(a_n)^{-1/n}$,

$$D(u) = 1 + \frac{a_1}{\sqrt{a_n}}u + \frac{a_2}{\sqrt{a_n^2}}u^2 + \dots + u^n \quad (2.29)$$

Factoring $D(u)$ into factors of 1st and 2nd orders,

$$D(u) = \begin{cases} (1 + c_{10}u) \prod_{k=1}^{(n-1)/2} (1 + c_{1k}u + u^2) \Leftarrow n \text{ odd} \\ \prod_{k=1}^{n/2} (1 + c_{1k}u + u^2) \Leftarrow n \text{ even} \end{cases} \quad (2.30)$$

Making the direct comparison with a normalized Butterworth polynomial of equal order (see appendix A 2), results in,

$$c_{10} = 1 \quad , \quad c_{1k} = 2 \sin \left[(2k-1) \frac{\pi}{2n} \right] \quad (2.31)$$

where k ranges from 1 to $n/2$ for n even, and $(n-1)/2$ for n odd. Equation (2.31) allows adjusting the

values of the circuit time constants in order to meet the maximal flat response constraint. However, as equations (2.30) and (2.31) reveal, systems of order larger than 1 must have characteristic equations factorable in quadratic terms, with a pair of complex poles each. This presents a fundamental design limitation: if the amplifier consists of a cascade of simple RC coupled stages without any feedback arrangements, then it is impossible to attain Butterworth maximum flatness design, since all the poles in the overall transfer function are necessary real. The amplifier will naturally present a flat frequency response, since the poles are all real. However, the achieved response will not be maximally flat in Butterworth sense, since there is no mechanism able to produce pairs of complex conjugate poles within the amplifier chain. The solution to this problem is to design each stage with quadratic characteristic equations. Two obvious solutions are: i) to use RLC coupled stages (generally speaking this resort to the design of two-terminal interstages or four-terminal interstages (which may possess higher order dynamics); or ii) to use second order feedback amplifiers in each interstage (possibly with real poles and adequate loop gain). In both solutions, the pairs of poles can be adequately adjusted to the sought Butterworth poles distribution, by a simple adjustment of the R , L and C values (or in the feedback case, adjusting the loop gain).

The GBW limitations for this Butterworth design procedure are the same than the GBW limitations of two-terminal or four-terminal interstage chain amplifiers (using feedback amplifiers may possess different kind of limitations that will be explored later). Thus there is a natural proximity between Butterworth design and frequency peaking techniques, since the interstages use naturally RLC elements in their conception.

2.6.2 Inductive Peaking

Figure 2.15 displays the most common forms of inductive peaking topologies. The first three circuits, representing series peaking topologies, allow larger bandwidth enhancement ratios. This

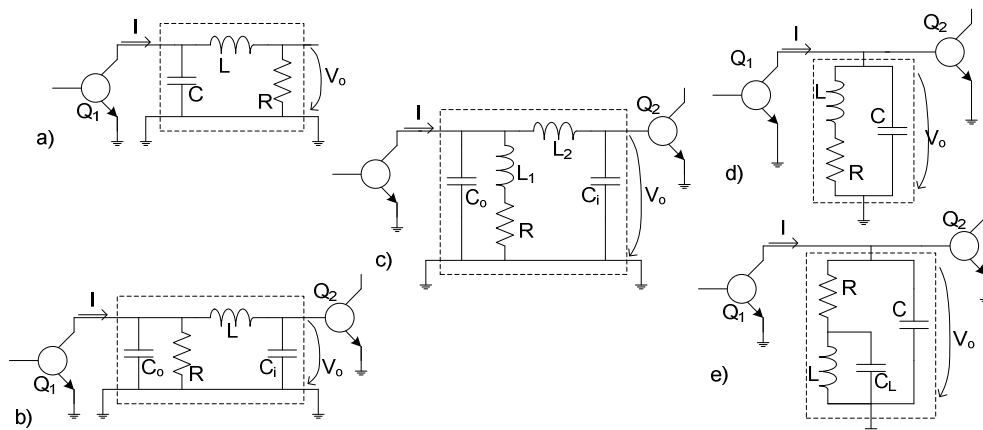


Figure 2.15 - Inductive peaking circuits: a) two pole series peaking, b) three pole series peaking, c) shunt-series peaking, d) two pole shunt peaking, e) three pole shunt peaking.

should not be surprising, since the series inductive element turns these topologies into examples of four-terminals interstages, with maximum BWER of 4,94. The last two circuits represent shunt peaking topologies. Clearly these are two-terminal interstages, having a maximum BWER of 2. The circuit of Figure 2.15c represents a hybrid between shunt peaked and series peaked, two inductive terms, one in series and other in parallel contribute to this fact.

The synthesis of inductive peaked circuits follows directly from the procedures discussed at the opening of this section. First a choice between maximum flat bandwidth and maximum flat delay should be considered. According to this choice the design procedure follows directly from the Butterworth or Bessel Polynomials, respectively. Other design constraints can be envisaged, for instance, it is common to specify circuits having the steepest possible transition band, using Chebyshev polynomials [251, 267]. However, since this work is devoted to GBW maximization, the strategy of choice is based on Butterworth design. For the sake of generality, it is possible to ensure that Butterworth design procedures are sufficiently close to their Bessel, Chebyshev counterparts (among others). The second stage of the design procedure is to compare the circuit's general transfer function with a normalized version designed using Butterworth polynomials. This is performed in a direct fashion for amplifiers having an all-pole low-pass type transfer function; since the polynomial to compare is the characteristic polynomial itself. When the transfer function has zeros (for instance, circuits of Figure 2.15c, d and e) the comparison is not direct. For these cases a more fundamental step is needed. The basic principle arises from a simple algebra theorem:

Theorem 2.1: *the necessary and sufficient conditions to ensure that two continuous and differentiable real valued functions of the same real variable, have some prescribed relation inside a defined interval, is that, the sought relation should be verified at the beginning of the interval and that the 1st derivative of the two functions, with respect to the common variable, have the same relation on the considered interval.*

This is just a generalization of the derivative properties of Butterworth polynomials. Butterworth polynomials satisfy this principle, all the first $n-1$ derivatives are 0 at the origin and also they all are strictly increasing functions of frequency. Considering the general transfer function defined as,

$$H(s) = \frac{N(s)}{D(s)} = \frac{b_0 + b_1s + \dots + b_ms^m}{a_0 + a_1s + \dots + a_ns^n} \quad (2.32)$$

with $n \geq m$. The corresponding frequency response magnitude is,

$$|H(j\omega)|^2 = \frac{(b_0 - b_2\omega^2 + b_4\omega^4 - \dots)^2 + \omega^2(b_1 - b_3\omega^2 + b_5\omega^4 - \dots)^2}{(a_0 - a_2\omega^2 + a_4\omega^4 - \dots)^2 + \omega^2(a_1 - a_3\omega^2 + a_5\omega^4 - \dots)^2} \quad (2.33)$$

Applying the strictly decreasing argument to (2.33), traduces in $|H(j\omega)|^2 \leq |H(0)|^2$ for all positive frequency values. Since $H(s)$ is assumed stable, $H(j\omega)$ has no singularities on the $j\omega$ axis, so it is possible to rearrange the terms of the same power. This leads to,

$$\frac{b_0^2 + (b_1^2 - 2b_0b_2)\omega^2 + (b_2^2 - 2b_1b_3 + 2b_0b_4)\omega^4 + \dots}{a_0^2 + (a_1^2 - 2a_0a_2)\omega^2 + (a_2^2 - 2a_1a_3 + 2a_0a_4)\omega^4 + \dots} \leq \frac{b_0^2}{a_0^2} \quad (2.34)$$

$$\dots + \left(b_k^2 + 2 \sum_{u=1}^k (-1)^u b_{k-u} b_{k+u} \right) \omega^{2k} + \dots + (b_{m-1}^2 - 2b_{m-2}b_m) \omega^{2m-2} + b_m^{2m} \omega^{2m}$$

$$\dots + \left(a_k^2 + 2 \sum_{u=1}^k (-1)^u a_{k-u} a_{k+u} \right) \omega^{2k} + \dots + (a_{n-1}^2 - 2a_{n-2}a_n) \omega^{2n-2} + a_n^{2n} \omega^{2n}$$

Comparing the terms of the same power results in a set of conditions,

$$\begin{aligned} a_0^2 (b_1^2 - 2b_0b_2) &\leq b_0^2 (a_1^2 - 2a_0a_2) \\ a_0^2 (b_2^2 - 2b_1b_3 + 2b_0b_4) &\leq b_0^2 (a_2^2 - 2a_1a_3 + 2a_0a_4) \\ &\dots \\ a_0^2 \left(b_k^2 + 2 \sum_{u=1}^k (-1)^u b_{k-u} b_{k+u} \right) &\leq b_0^2 \left(a_k^2 + 2 \sum_{u=1}^k (-1)^u a_{k-u} a_{k+u} \right) \\ &\dots \\ a_0^2 b_m^{2m} &\leq b_0^2 \left(a_m^2 + 2 \sum_{u=1}^m (-1)^u a_{m-u} a_{m+u} \right) \\ 0 &\leq a_{m+1}^2 + 2 \sum_{u=1}^{m+1} (-1)^u a_{m+1-u} a_{m+1+u} \\ &\dots \\ 0 &\leq a_{n-1}^2 - 2a_{n-2}a_n \end{aligned} \quad (2.35)$$

These conditions translate into the general rule that, for a rational transfer function to fulfill maximum flatness, it is necessary and sufficient that all the terms having the same power, between numerator and denominator of the corresponding squared magnitude frequency response must be equal⁸ [50]. Since in general $n \geq m$ all the remaining denominator powers have positive or zero coefficients. In general the verification procedure of (2.35) is not as complicated as suggested. For the majority of practical cases the orders of numerator and denominator polynomials are modest; for instance, the worst case in Figure 2.15 is the shunt-series topology, for which the numerator and denominator polynomials are of 4th and 1st orders, respectively.

The design example of a two-pole series peaking circuit and a two-pole shunt peaking circuit will be briefly discussed, as an illustration of the design procedure. Considering first the two-pole

⁸ Alternatively, it is possible to apply a simple Butterworth design procedure, taking $a_o N(s) - b_o D(s)$ instead.

series peaking circuit of Figure 2.15a the general form of the gain of this circuit is proportional to the transfer impedance of the peaking network, that is,

$$Z_T(s) = \frac{V_o(s)}{I(s)} = \frac{R}{1 + sRC + s^2LC} \quad (2.36)$$

Using the normalizing transformations $\omega_r = 1/RC$ and $L = mR^2C$ (note that ω_r represents the cut-off frequency of the non-peaked version of the same circuit), and comparing the result with a normalized Butterworth polynomial of second order with cut-off frequency ω_c ,

$$1 + \left(\frac{s}{\omega_r}\right) + m\left(\frac{s}{\omega_r}\right)^2 = 1 + \sqrt{2}\left(\frac{s}{\omega_c}\right) + \left(\frac{s}{\omega_c}\right)^2 \quad (2.37)$$

from which m must be set to $1/2$. This also reveals that the BWER achieved is simply $\omega_c/\omega_r = 2^{1/2}$, that represents only 28,5% of the maximum available BWER of a four-terminal interstage.

Considering now the shunt-peaked circuit (Figure 2.15d), the gain of the first stage is proportional to the driving point impedance of the interstage (since for this case the interstage has a two-terminal topology rather than a four-terminal has in the previous case), that can be written as,

$$Z(s) = \frac{V_o(s)}{I(s)} = \frac{R + sL}{1 + sRC + s^2LC} \quad (2.38)$$

This is an example of a rational transfer function of the type previously discussed. Performing the same transformations as for the above example, equation (2.38) becomes,

$$Z(s) = \frac{R \left[1 + m \left(\frac{s}{\omega_r} \right) \right]}{1 + \left(\frac{s}{\omega_r} \right) + m \left(\frac{s}{\omega_r} \right)^2} \quad (2.39)$$

Identifying in (2.39) the terms of a_k and b_k , the conditions in (2.35) can be directly applied,

$$m^2 \leq 1 - 2m \quad (2.40)$$

This condition has to be fulfilled in an equality sense, thus leading to $m = 2^{1/2} - 1$. For this case the BWER achieved does not follow as directly as in the previous case and it is necessary to evaluate the cut-off frequency. Assuming the usual $3dB$ definition, this corresponds to $\omega_c = 1.72\omega_r$, corresponding to a BWER of 1.72 and to a 86.1% usage of the available BWER of a two-terminal interstage. In simple terms, shunt peaking is more efficient than series peaking because of the effective usage of the available GBW. However, comparing the performance in terms of BWER, series peaking circuits have larger bandwidth for the same capacitance-resistance values.

Table 2.1 presents the normalized transfer functions for all the circuits depicted on Figure 2.15: the second column shows the normalized transfer function proper, where the frequency variable u is given by $u = s/\omega_r$, with $\omega_r = 1/RC$; the third column shows the normalizing transformations (for the

Table 2.1 - Normalized transfer functions for the circuits of Figure 2.15.

	$H_n(u)$	Transformations	MFD	BWER/ η
Fig 2.15a	$\frac{1}{1+u+m_1u^2}$	$L = m_1R^2C$	$m_1 = 0.5$	$\sqrt{2}/28.6\%$
Fig 2.15b	$\frac{1}{1+u+m_1m_2u^2+m_1m_2(1-m_2)u^3}$	$L = m_1R^2(C_i + C_o)$ $m_2 = C_i/(C_i + C_o)$	$m_1 = 2/3$ $m_2 = 3/4$	$2/40.5\%$
Fig 2.15d	$\frac{1+m_1u}{1+u+m_1u^2}$	$L = m_1R^2C$	$m_1 = \sqrt{2} - 1$	$1.72/86\%$
Fig 2.15e	$\frac{1+m_1u+m_1m_2u^2}{1+u+m_1(1+m_2)u^2+m_1m_2u^3}$	$L = m_1R^2C$ $C_L = m_2C$	$m_1 = 0.414$ $m_2 = 0.354$	$1.84/92\%$
Fig 2.15c	$\frac{1+m_1u}{1+u+(m_1+m_2m_3)u^2+m_2m_3(1-m_3)u^3+m_1m_2m_3(1-m_3)u^4}$	$m_1 = L_1$ $m_2 = L_2$ $m_3 = C_i$	$m_1 = 0.1464$ $m_2 = 0.6036$ $m_3 = 0.5858$	$2.61/52.8\%$

case of Figure 2.15c these transformations are slightly different, the cut-off frequency $\omega_c = 1/R(C_i + C_o)$ is assumed unitary, with both $R=1$ and $C_i + C_o=1$; the fourth column presents the conditions to meet maximum flatness; finally, the last column shows the correspondent BWER under the maximal flatness constraint and the efficiency of the design (using $\eta = \text{BWER}/\text{BWER}_T$, where BWER_T is the available BWER of a two-terminal or four-terminal interstage, accordingly). It is instructive to observe some aspects of these normalized transfer functions.

Concerning the two-terminal interstage topologies (Figure 2.15d and Figure 2.15e), the correspondent transfer functions have the form of realizable driving point impedances, with one zero less than the number of poles. This argument fixes the total phase variation produced in $\pi/2$ radians. Increasing the complexity of the interstage is an expedite way to further enhance the achieved BWER. However, it is noticeable that with these structures almost all of the available BWER is effectively used.

The remaining circuits have a four-terminal interstage topology, thus the available BWER is larger than for the previous cases. The moderate BWER achievements of these circuits clearly show that there is much bandwidth to explore increasing the complexity. Figure 2.16 shows the magnitude and phase of the frequency response for the above normalized transfer functions, using the maximal flat constraint as reference. The effectiveness of the above solutions for bandwidth enhancement is evident.

The cost of maximum flatness is also relevant; enhancing the bandwidth by a correct allocation of the poles also has effect on the phase response. As the bandwidth increases, the phase response becomes less and less linear, meaning that the delay becomes more and more frequency dependent. This effect is even more pronounced on four-terminal structures, since the total phase variation is not restricted to 90° variation. A solution to this problem is to tune the circuit parameters in order to

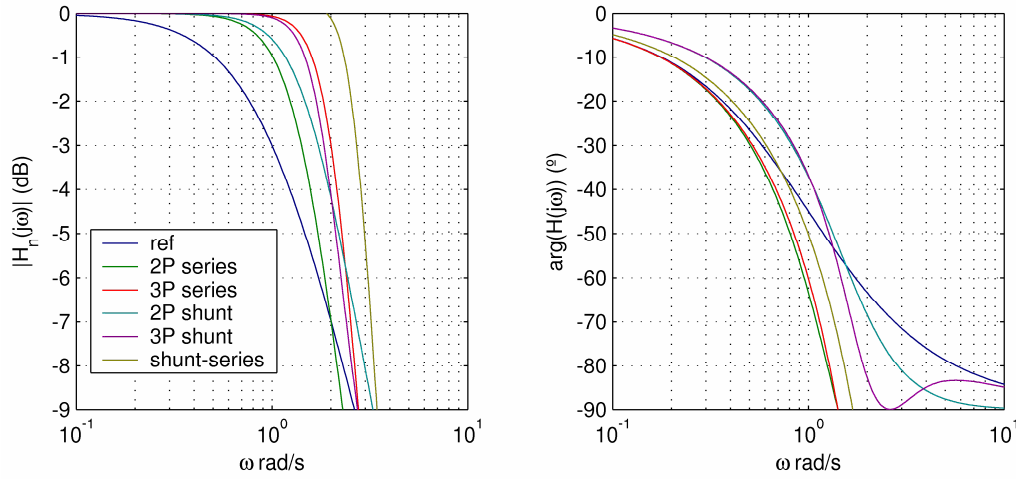


Figure 2.16 - Frequency response under maximally flat constraints of the circuits of Figure 2.15.

achieve flat delay instead of flat gain.

2.6.3 T-Coil Peaking

Several reported results using circuit topologies similar to those of Figure 2.15 confirm the above results [51-53, 55-63]. Larger BWER values are possible increasing the complexity of the interstages, good examples of these are several forms of interstage design based on T-Coils [57, 58, 60, 286]. T-Coil interstages were broadly used in distributed amplifiers, as a means of transmission line modeling (see for instance [42, 43]). The basic T-Coil circuit is depicted on Figure 2.17. Generally speaking, the T-Coil circuit is based on a center tapped inductance, bridged by the capacitance C_b , consisting ideally on the inductance self-capacitance. Another important factor is the coupling between the two inductors, k ; the relation between k and C_b is a key design parameter for these kind of circuits. With a shunting capacitance C , the T-Coil circuit resembles an infinitesimal element of a transmission line, thus justifying the affinity with distributed amplifiers. Considering the T-Coil part of a general interstage, it can be assumed that one of its ports is terminated on a load resistance R (the output port). Neglecting C_b , an immediate relation that must be fulfilled in order to have constant input impedance (on the opposite port), is that $R=(L/C)^{1/2}$; satisfying this restriction, follows that the input impedance is given directly by R .

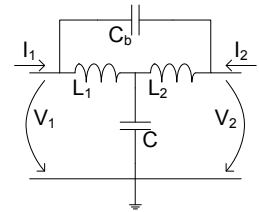


Figure 2.17 - Basic T-Coil circuit.

Several forms of four-terminal interstages including T-Coils can be designed. The most common ones are represented in Figure 2.18. The simplest form is the two-pole T-Coil interstage. For this case, if the output is taken from the loading resistor, the overall transfer function is of all-pass type. However, in order to explore frequency peaking phenomena, the output has to be taken from the capacitance C instead. Considering the center tapped inductance of Figure 2.18a with coupling

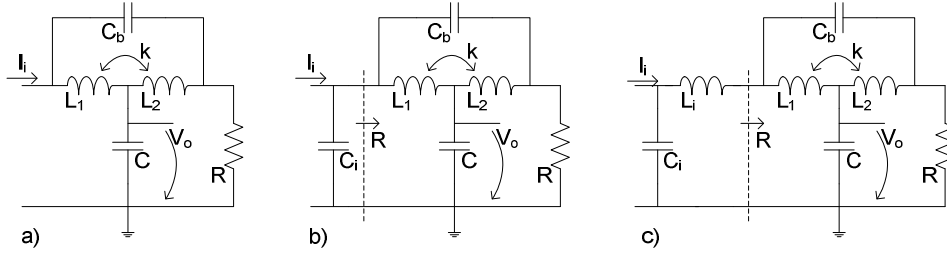


Figure 2.18 - T-Coil interstages: a) two-pole T-Coil; b) three-pole T-Coil; c) four-pole L+T.

factor k , the effective inductance is given by $L=L_1+L_2+2L_M$, and $L_M=k(L_1L_2)^{1/2}$; using this transformation it is possible show that for,

$$\begin{aligned} L &= R^2 C \\ L_M &= R^2 \left(\frac{C}{4} - C_b \right) \end{aligned} \quad (2.41)$$

the overall transfer function becomes,

$$H(s) = \frac{V_o(s)}{I_i(s)} = \frac{R}{1 + sRC/2 + s^2 R^2 CC_b} \quad (2.42)$$

Assuming a maximally flat design, the denominator of (2.42) must be compared to a normalized second order Butterworth polynomial, resulting on the following conditions: $C_b=C/8$, $L_M=R^2C/8$ and $k=0.33$, where the value of k holds for $L_1=L_2=L/2-L_M$ (with $L=R^2C$). The correspondent BWER of this circuit is $8^{1/2}$ (obtained by comparison against $\omega_c=1/RC$), with an associated efficiency of 57.3%. It is possible to further enhance the BWER factor by simply allowing some gain variation, however the phase characteristic becomes seriously compromised.

A limiting factor for T-Coil designs is the necessity of realizing sufficiently accurate inductances with precise coupling factors, especially hard to do on integrated circuit design. Despite this stringent limitation, the usage of T-Coils in modern CMOS integrated circuits operating at 40GHz has been demonstrated [57, 58, 60]. Since it can be difficult to control the coupling factor, one possibility is to design the inductances without coupling. This leads to $C_b=C/4$ and both L_M and k equal to 0. Unfortunately, this leads to large overshoot values on the step response ($\sim 16.3\%$).

A more efficient design approach is to use the three-pole T-Coil circuit (Figure 2.18b). For this case, using the same maximally flat constraint, naturally results in a coupling factor of 0. Assuming that the T-Coil circuit is designed with constant input impedance equal to R , the overall transfer function consists of the two poles of (2.42) plus the contribution of a third input pole due to C_i .

$$H(s) = \frac{V_o(s)}{I_i(s)} = \frac{R}{(1 + sRC_i)(1 + sRC/2 + s^2 R^2 CC_b)} \quad (2.43)$$

This simplified analysis results from the fact that the T-Coil network can be made reciprocal and thus it has no effect on the input pole. Comparing with an adequate third order Butterworth

polynomial, and adding the previous inductance relations, the conditions for maximal flat frequency response are; $C_i=C/2$, $C_b=C/4$ and $k=0$. The correspondent BWER is 3 and the associated efficiency is 60.7% (obtained by comparison against $\omega_r=1/(C+C_i)$).

Better arrangements can be obtained using the above considerations. If the T-Coil is designed with constant input impedance then it is possible to use an L section cascaded at the input of the T-Coil network, resulting in the L+T circuit of Figure 2.18c. The overall transfer function is,

$$H(s) = \frac{V_o(s)}{I_i(s)} = \frac{1}{1 + sRC_i + s^2mR^2C_i^2} \frac{R}{1 + sRC/2 + s^2R^2CC_b} \quad (2.44)$$

where $L_i=mR^2C_i$. Defining r_1 and r_2 as $2\sin(\pi/8)$ and $2\sin(3\pi/8)$, respectively and comparing the denominator of (2.44) with a normalized fourth order Butterworth polynomial, results in the following relations; $C=4r_2^2C_b$, $C_i=2r_1r_2C_b$, $m=r_1^{-2}$ and $k=(r_2^2-1)/(r_2^2+1)$. This corresponds to an amazing BWER, given by $2r_2+r_1 \approx 4.46$, with an associated efficiency of 90.3%. The major limitation of this topology is the necessity of the large values of the coupling factor k .

The intention of this simplified analysis was to show how the GBW limit of a four-terminal interstage can be effectively approached using matching networks with moderate complexity. A more detailed analysis of T-Coil circuits can be found in [286]. A close inspection of the achieved BWER values for all the inductive peaking circuits also reveals that there is no major justification to further increase the complexity, since there is not much to gain. Furthermore, increasing the complexity of the matching networks also adds more design variables to the process, becoming more and more difficult to synthesize and explore its benefits. A theoretical justification of these facts was addressed by Fano [25, 246], in his work about the theoretical limitations of arbitrary matching.

2.6.4 Capacitive Peaking

Capacitive peaking strategies borrow their name from a parallel with inductive peaking. Similar to inductive peaking, it is possible to explore in some special circuits significant bandwidth improvements, when some adequately placed capacitance changes its value. The idea of capacitive peaking is not intuitive; in general capacitances act as bandwidth restricting elements, so the idea of adding circuit capacitances in order to increase bandwidth is at least unnatural. Inductive peaking is based on creating resonant frequencies by the addition of inductive elements to the circuit. It is also possible to think in resonant frequencies in circuits without these inductive elements; for instance, feedback amplifiers having more than one pole may, under some special conditions exhibit frequency peaking. Another example is the usage of circuits employing inductance simulators (circuit arrangements able to reproduce inductive frequency behaviors). The first example is simpler

than the second, as in general inductance simulators results in added complexity which in its turn further restricts the overall bandwidth.

In feedback amplifiers, it is possible to enhance bandwidth by a careful adjustment of the system poles (and/or zeros). One such example is the presence of a load capacitance at the output port of the amplifier. This capacitance can act as an element defining the output pole. Depending on feedback, this capacitance alone can turn the amplifier unstable.

This observation has driven many compensation techniques suitable for feedback amplifiers based on two-stage operational amplifiers [64-69, 73, 74]. However, for a particular closed-loop gain, this output capacitance can be adjusted in order to increase bandwidth. This is just the motive that led Chien *et al* to propose the capacitive peaking design strategy in 1999 [70-72]. The

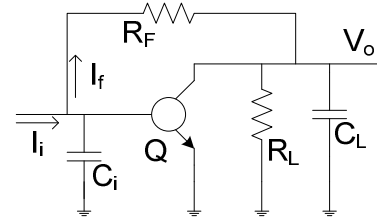


Figure 2.19 - Capacitive peaking concept.

concept can be briefly explained using a transimpedance amplifier⁹ based in only one transistor, as depicted on Figure 2.19. The closed-loop transimpedance gain is derived using the open-loop gain; this can be obtained breaking the feedback loop – that is the feedback resistance R_F , loads in parallel both input and output ports. The open loop gain follows directly,

$$\begin{aligned} Z_o(s) &= -\frac{g_m R_F (R_F // R_L)}{(1 + s R_F C_i)(1 + s (R_F // R_L) C_L)} \\ &\approx -\frac{g_m R_F^2}{(1 + s R_F C_i)(1 + s R_F C_L)} \end{aligned} \quad (2.45)$$

where g_m represents the transistor's transconductance and the approximation holds if $R_L \gg R_F$. The closed-loop gain is now evaluated knowing that the feedback factor is simply $\beta = -1/R_F$,

$$\begin{aligned} Z_F(s) &= \frac{Z_o(s)}{1 + \beta(s) Z_o(s)} \\ &\approx -\frac{g_m R_F^2}{1 + g_m R_F} \frac{1}{1 + s \frac{R_F (C_i + C_L)}{1 + g_m R_F} + s^2 \frac{R_F^2 C_i C_L}{1 + g_m R_F}} \end{aligned} \quad (2.46)$$

Using a maximal flat Butterworth design approach, the maximum BWER that can be achieved with this arrangement is a factor of $2^{1/2}$. Defining $\tau_i = R_F C_i$, $\tau_o = R_F C_L$ and $r = \tau_o / \tau_i$, the conditions for maximal flatness are given as,

⁹ Capacitive peaking is often related with feedback amplifiers. In fact, it is the feedback dynamics that make possible the exploitation of some special capacitances as bandwidth enhancement elements. Nevertheless, due to its acquired importance in amplifier design, it seemed reasonable to present capacitive peaking as another form of frequency peaking, rather than, a generic feedback design concept.

$$\sqrt{1+g_m R_F} = \frac{1+r}{\sqrt{2r}} \quad , \quad \omega_n = \frac{1+r}{r\tau_i\sqrt{2}} \quad (2.47)$$

Equation (2.47) suggests that the cut-off frequency, ω_n , may become infinite if $C_L=0$ ($r=0$). This results because the loop-gain would become infinite for this situation and the system would degenerate into a one pole transfer function. If this cut-off frequency is compared against the cut-off frequency of a reference amplifier having the same closed-loop gain without C_L (given as $(1+g_m R_F)/\tau_i$), the resulting BWER is given by,

$$BWER = \frac{\omega_n}{\omega_r} = \frac{\sqrt{2}}{1+r} \quad (2.48)$$

Equation (2.48) shows that the achievable maximum BWER under maximal flatness constraint decreases with increasing C_L . In fact the $2^{1/2}$ factor can only be achieved for small values of r .

Equating the real BWER without the maximal flatness restriction reveals other aspects of the technique. Figure 2.20 shows the BWER behavior of this technique. The bold red line shows the maximum BWER under maximal flatness according to (2.48). The blue lines reveal the peaking phenomena associated with the ratio C_L/C_i for various loop-gain values. As can be seen, it is possible to increase C_L and enhance bandwidth until the maximum flatness limit is not reached. From this point onwards, bandwidth decreases with increasing C_L values.

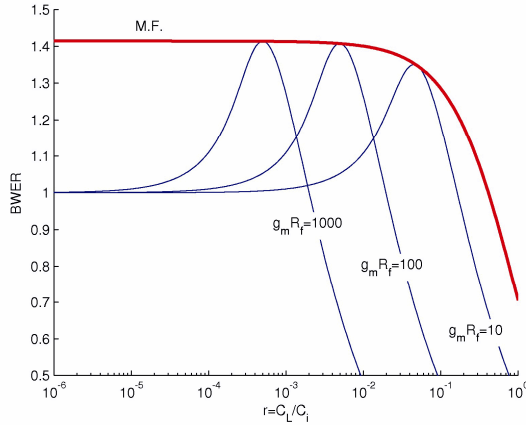


Figure 2.20 - BWER using capacitive peaking.

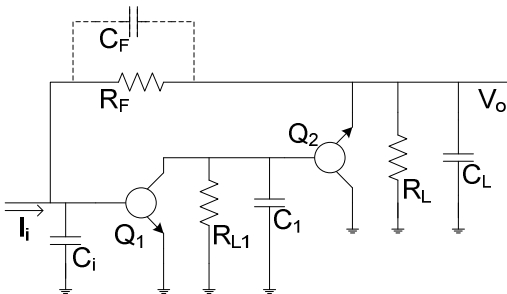


Figure 2.21 – Enhanced circuit version Figure 2.19.

A real implementation of this concept has to consider several aspects, namely: i) the approximation in (2.45) may not hold if R_L is comparable to R_F ; ii) the implementation of the feedback resistance is not free from parasitic capacitances. The first problem can be easily solved adding an output buffer to the circuit of Figure 2.19 as represented on Figure 2.21. For this case the output resistance of the second stage can be made arbitrarily low; resulting in R_{o2} being smaller than R_F , meaning that there are no loading effects caused by the output stage. Another consequence of adding an output buffer is the addition of more poles into the amplifier's transfer function. This can be regarded as a benefit, since it also allows larger BWER values to be achieved.

The influence of the parasitic capacitance due to the feedback resistance, C_F , traduces into a RHP zero, which also allow larger BWER values, once the adequate optimization strategies (aiming for maximally flat frequency response) have been applied [96, 106].

Other capacitive peaking strategies have been reported in [64, 66]. These examples aim to cancel parasitic capacitances in active devices. It is commonly known that the capacitance c_z of an active device (c_μ for a BJT and c_{gd} for a FET) poses severe restrictions on bandwidth achievements of CX amplifying stages. This bandwidth limitation appears as result of Miller multiplication of c_z at the input port of such amplifiers. In [64], Mataya *et al* proposes a strategy to cancel the capacitance c_z in differential structures employing two parasitic transistors. On [66], Wakimoto *et al* apply a different kind of cancellation based on positive feedback with the same aim.

2.7 Feedback Amplifiers

Feedback amplifiers were originally proposed in a 1937's patent by Black [75]. In its original conception, a feedback amplifier comprised two distinct signal paths: the forward path, responsible for signal amplification; and a feedback path, used to convey a sample of the output signal back to the input of the amplifier. In his patent, Black reported several possible arrangements to synthesize feedback amplifiers. The topologies used by Black have some resemblance to the actual feedback configurations; for instance the term voltage-voltage feedback was applied for both series-series and series-shunt feedback. Although the nomenclature and the analysis methods have evolved since then, the basic properties of feedback amplifiers remained the same: enhanced immunity to component variations, enhanced linearity, simple and efficient method to achieve gain control, and enhanced input and output impedance characteristics. Black addressed also the problem of stability of feedback amplifiers, but bandwidth and general GBW limitations were left uncovered. Since then many refinements to feedback amplifier design were proposed (see for instance [76, 84, 102, 247, 252, 264, 267, 277, 280]).

2.7.1 GBW Properties of Feedback Amplifiers

Figure 2.22 displays a feedback amplifier comprising a forward amplifier, $A(s)$, and a feedback path with transmission gain $\beta(s)$. The sign convention on the adding element at the input of the amplifier identifies the feedback type: negative or degenerative if the feedback signal x_f subtracts itself to the input signal x_i ; and positive or regenerative if the feedback signal reinforces the input signal. Obviously, the feedback can assume negative or positive behaviors depending on the phase rotation around the

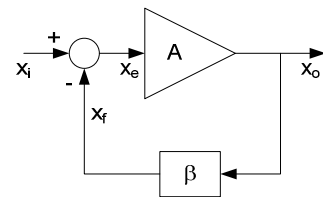


Figure 2.22 - Feedback amplifier.

feedback loop. This means that amplifiers having more than two poles or reactive elements on the feedback network, are able to produce a total phase rotation exceeding 180° and thus, may exhibit both positive and negative feedback types for different frequency bands.

An important property of feedback amplifiers is their GBW constancy. A simple demonstration of the GBW constancy property of feedback amplifiers can be devised using a simple first order amplifier as reference. Assuming that: i) the pole of the open-loop amplifier is at ω_o rad/s, and its DC gain is A_o ; ii) the feedback factor β is constant (this guarantees frequency independent feedback); iii) there is no loading effects both at the input and output of the amplifier. With these assumptions, the GBW of the open-loop amplifier is simply $A_o\omega_o$. The closed-loop transfer function is given as,

$$A_{CL}(s) = \frac{A_o}{1 + \beta A_o} \frac{1}{1 + \frac{s}{\omega_o(1 + \beta A_o)}} \quad (2.49)$$

for which the first order model still applies. The GBW of the closed-loop amplifier is thus,

$$\begin{aligned} GBW_{CL} &= \frac{A_o}{1 + \beta A_o} \omega_o (1 + \beta A_o) \\ &= A_o \omega_o \end{aligned} \quad (2.50)$$

Equation (2.50) proofs that the GBW of an idealized first order feedback amplifier is independent of the feedback. This is a generally accepted result on feedback amplifiers, nevertheless, its validity is very limited. From the GBW view point, much can be done on feedback amplifiers having more than one pole. For instance, it is possible to find many examples of feedback amplifiers having an open-loop dominant pole behavior and a closed-loop GBW that surpasses the open-loop reference. This is the case of feedback amplifiers having at least second order dynamics designed to present maximally flat frequency response [279, 284].

2.7.2 *Maximally Flat Design*

It is possible to design feedback amplifiers with maximally flat frequency response for a particular set of gains A_o and β . These would also allow to explore the theoretical GBW limitations of feedback amplifiers. Assuming for the amplifier of Figure 2.22 that $A(s)$ is given by $A_o/D(s)$ (with $D(0)=1$) and $G=\beta A_o$, the maximum achievable theoretical GBW can be found if, for some design conditions the closed-loop amplifier exhibits a maximally flat frequency response (it is traditional to use this design constraint for amplifiers having 2 poles). Starting with the characteristic equation,

$$D(s) + G = (1 + G) \left(1 + \frac{a_1}{1+G} s + \frac{a_2}{1+G} s^2 + \dots + \frac{a_n}{1+G} s^n \right) \quad (2.51)$$

Now, assuming that for an optimum value of G , it is possible to have the closed-loop poles placed on a Butterworth circle of radius ω_o (thus assuring the maximally flat constraint), equation (2.51) can be compared with,

$$D(s) + G = (1 + G) \left(1 + b_1(n) \frac{s}{\omega_o} + b_2(n) \frac{s^2}{\omega_o^2} + \dots + \frac{s^n}{\omega_o^n} \right) \quad (2.52)$$

where $b_k(n)$ are the coefficients for an n^{th} order normalized Butterworth polynomial. The resulting bandwidth is related to the loop gain by,

$$\omega_o = b_1(n)(1 + G)/a_1 \quad (2.53)$$

For a system having one pole at ω_l rad/s, a_1 can be made as close to $1/\omega_l$ as desired. Knowing also that $b_1(n)$ is given by $1/\sin(\pi/2n)$, the maximum possible closed-loop GBW for a n^{th} order feedback amplifier is,

$$GBW_{CL}(n) = \frac{A_o}{1 + G} \omega_o \approx \frac{A_o \omega_l}{\sin(\pi/2n)} \quad (2.54)$$

Equation (2.54) shows that there is an inherent GBW benefit in increasing the number of poles in a feedback system. Furthermore, it implies that there is no theoretical GBW limitation for feedback amplifiers having an infinite number of poles. It is interesting to notice that since the bandwidth increases with the number of poles, the cut-off frequency might become larger than the transition frequency. This obviously triggers some sort of GBW limitation. According to Bode results, the transition frequency plays a key role on the maximum achievable GBW of both distributed and non-distributed amplifiers, it seems logical to assume that the same sort of limitation must exist for feedback amplifiers. In conclusion, a feedback amplifier having an infinite number of poles in the sense of (2.54) is only a theoretical abstraction.

Increasing the number of poles in a feedback amplifier has its drawbacks, namely:

- Increasing the order of the characteristic equation lessens the stability margins (for a fixed bandwidth, the necessary loop gain to achieve maximum flatness, decreases with $\sin(\pi/2n)$);
- The open-loop poles have prescribed possible trajectories, if the maximum in (2.54) is to be strictly achieved. These trajectories can be found using a backward root-locus algorithm for all possible values of G , departing from the Butterworth poles. Figure 2.23 show these trajectories for 3rd and 4th order cases, using a normalized cut-off frequency of 1 rad/s (different cut-off frequencies produce scaled versions of these pictures);
- From the practical point of view it is very difficult to realize the correct pole allocations for

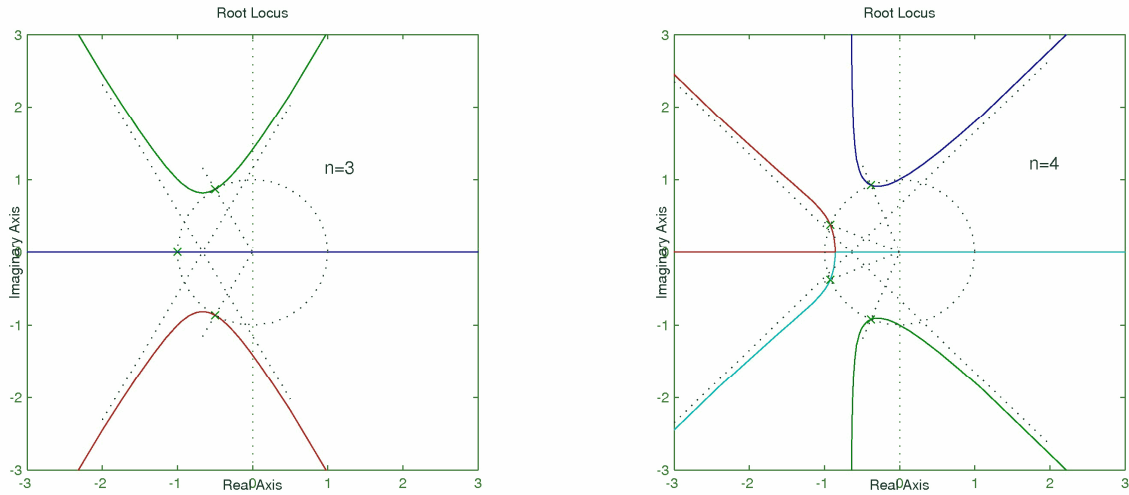


Figure 2.23 - Root locus trajectories for maximally flat feedback amplifiers having 3rd and 4th order dynamics.

systems of 3rd order and more. The presence of at least one pair of complex conjugate poles (for $n > 2$), turn system design difficult to accomplish. Two possibilities of having the presence of 2nd order terms inside a transfer function are to use inductances or local feedback loops. The first possibility resorts to the design of inductive peaked interstages having prescribed pole allocations. In view of what was previously discussed, these interstages must be necessarily four-terminal interstages, thus satisfying the requirement of having only poles. Adding zeros to the overall transfer function would result in more complex design procedures. The second possibility is to use local feedback loops to place the poles adequately. The drawback of this solution is the natural restriction on gain, imposed by the necessary loop-gain. This would become worse for large n .

Another source of restrictions is obviously the increased complexity. Needless to say, that increasing the circuit complexity also increases the sources of error; many components imply too many tolerances to account! This justifies the common practice of considering the maximal flat design for feedback amplifiers having only 2nd or 3rd order dynamics [85, 90, 279, 280, 284, 286]. Generally this simplification is taken one step further, by assuming that the amplifier is composed of an RC active chain having only real poles, and that from these poles only the first two or three contribute significantly to the amplifier's frequency response.

2.7.3 Variable Feedback

Until now it was assumed that the feedback factor was frequency independent. However gain improvements may be achieved if this is not so. Figure 2.24 represents the frequency response of a feedback amplifier having an open-loop gain of 60dB and a cut-off frequency of 1 rad/s. The left side of Figure 2.24 represents the constant feedback case; assuming a feedback factor of -40dB, the

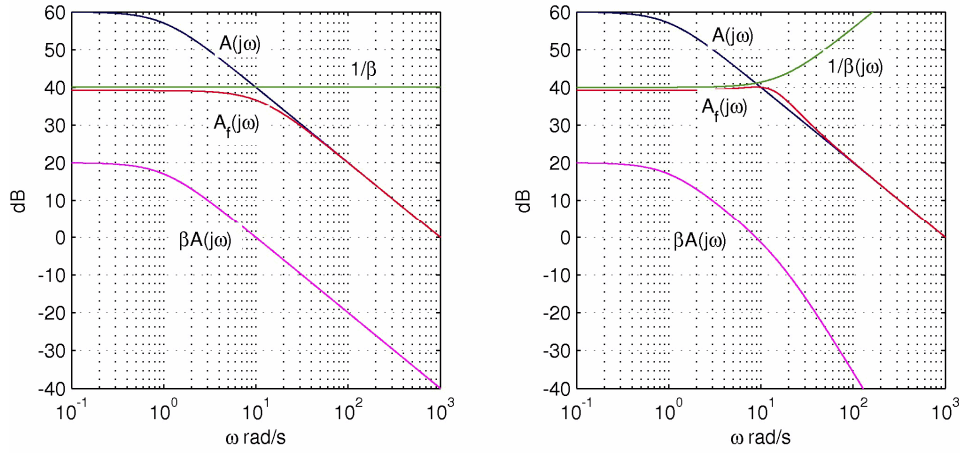


Figure 2.24 - Feedback amplifier's frequency response: constant feedback (left side), variable feedback (right side). closed-loop gain reduces to almost 40 dB with cut-off frequency increased to 11 rad/s . It is possible to further increase the cut-off region allowing the feedback factor to change its value with frequency. An immediate example of this principle is depicted on the right side of Figure 2.24. For this example the feedback factor has one pole near the cut-off of the closed-loop amplifier. As Figure 2.24 shows, the feedback factor, $1/\beta(j\omega)$ increases with frequency allowing for gain reduction compensation due to the feedback. Significant bandwidth improvements can be achieved if this effect is tuned to the cut-off region. Considering β_o and ω_β the zero frequency value and the cut-off frequency of β factor, the closed-loop frequency response can be written as,

$$\begin{aligned}
 A_f(s) &= \frac{A(s)}{1 + \beta(s)A(s)} = \\
 &= \frac{A_o}{1 + \beta_o A_o} \frac{1 + \frac{s}{\omega_\beta}}{1 + \frac{s}{\beta_o A_o} \frac{\omega_o + \omega_\beta}{\omega_o \omega_\beta} + \frac{s^2}{(1 + \beta_o A_o) \omega_o \omega_\beta}}
 \end{aligned} \tag{2.55}$$

Maximal flat frequency response can be achieved using the design conditions developed in (2.35), and this results in the following value of ω_β ,

$$\frac{\omega_\beta}{\omega_c} = \left(1 - \frac{1}{(1 + \beta_o A_o)} + \sqrt{2} \sqrt{1 - \frac{1}{(1 + \beta_o A_o)}} \right) \tag{2.56}$$

Where ω_c is the cut-off frequency of a reference amplifier with constant β . For large values of $1 + \beta_o A_o$, (2.56) approaches $1 + 2^{1/2}$. The cut-off frequency of the compensated amplifier can achieve a maximum bandwidth improvement of 1.722 (the same improvement achieved with a two-pole shunt peaking interstage, since the transfer function has a similar form). Arrangements displaying such feedback dependence are similar to the case discussed on the previous section relating to Figure 2.21, where the feedback loop is composed of a resistance plus its parasitic capacitance.

Several variations of this variable- β technique have been explored in the past. Using capacitive elements in the feedback path is a simple technique to turn the feedback frequency dependent; this is the case of the works reported in [84, 85, 90, 91, 96, 108]. The authors of [84, 96] explored the usage of phantom zeros, capacitive feedback and another forms of variable feedback, as means of bandwidth optimization.

2.7.3.1 Capacitive Feedback

The concept of capacitive feedback was originally discussed by Vadipour in [96]. In This original contribution, capacitive feedback is described as a bandwidth enhancement technique suitable for differential amplifiers. Nevertheless, the concept is perfectly general and can be applied to other feedback configurations. Refinements to this technique were proposed by Centurelli *et al* in [106].

Capacitive feedback can take several forms, depending on the feedback configuration. The important aspects of capacitive feedback are: i) the possibility of bandwidth enhancement; ii) the fact that it does not affect the DC gain, as opposed to normal resistive feedback. A simple illustration of this concept is represented in Figure 2.25, where a transimpedance amplifier employing shunt-shunt capacitive feedback is used. Considering that the open-loop amplifiers is adequately represented by a first order transfer function with DC gain $Z_o = A_v r_i$ (where, A_v and r_i represent the voltage gain and the input impedance, respectively), and open-loop pole $1/\tau_i$ (where τ_i represents the input time constant due to feedback loading and amplifier input circuitry), the closed-loop transfer function becomes,

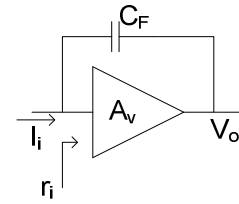


Figure 2.25 - Shunt-shunt capacitive feedback.

$$Z_{CL}(s) = \frac{\pm Z_o}{1 + s(\tau_i \mp Z_o C_F)} \quad (2.57)$$

Equation (2.57) can have different interpretations depending on the signal of Z_o . If Z_o is negative, the feedback is also negative and the closed-loop cut-off frequency is reduced, since τ_i is augmented by the factor $Z_o C_F$. This is a simple illustration of the Miller effect; the feedback capacitance appears multiplied by the gain on the closed-loop transfer function. However, a different result happens when Z_o is positive, for this case the feedback is positive and the cut-off frequency can be increased without bound, since $Z_o C_F$ reduces the effect of τ_i .

In real circuits the above simplified analysis can be seriously compromised. Two factors contribute to prevent such ideal behavior, namely: i) in general, the open-loop transfer function has more than one pole - at least two poles are imposed by loading effects originated by the feedback

capacitance; ii) stability restrictions prevent the “boundless” bandwidth increase – RHP poles may appear if the dominant time constant becomes negative. Considering these factors, the closed-loop transfer function for a 2nd order amplifier is given by,

$$Z_{CL}(s) = \frac{Z_o}{1 + s(\tau_i + \tau_o + Z_o C_F) + s^2 \tau_i \tau_o} \quad (2.58)$$

where τ_i and τ_o represent the two poles (possibly due to the input and output of the amplifier's circuitry). Assuming maximum flat frequency response operation, the normalized denominator of (2.58) must match a second order Butterworth polynomial, resulting in the following relation,

$$C_F = \frac{1}{\omega_c Z_o} \left[\sqrt{2} - \sqrt{\frac{\tau_i}{\tau_o}} \left(1 + \frac{\tau_o}{\tau_i} \right) \right] \quad (2.59)$$

where $\omega_c = 1/\sqrt{\tau_i \tau_o}$ is the achieved cut-off frequency given as the geometric mean of the frequencies of the open-loop poles. Equation (2.59) reveals that both negative and positive feedback types can be employed, depending on the open-loop poles. Assuming that $\tau_i \gg \tau_o$, the negative part inside brackets of (2.59) might become larger than $\sqrt{2}$, implying that Z_o has to be negative, leading to negative feedback. For this case, the BWER is proportional to $\sqrt{\tau_i/\tau_o}$. The opposite situation leads directly to positive feedback, and the resulting BWER is proportional to $\sqrt{\tau_o/\tau_i}$.

2.7.3.2 RHP Zeros

Miller effect offers another form of capacitive feedback. Amplifier circuits having negative voltage gain between two nodes may exhibit Miller effects. If these nodes are connected through a capacitance, Millers theorem says that the equivalent capacitance seen at the input node appears multiplied by the voltage gain. This capacitance multiplication effect represents a major bandwidth limitation in amplifier circuits.

Traditional amplifier design techniques often use Miller's approximation to model the amplifier dynamics. In general, this approach conducts to approximated results that are sufficiently accurate and easy to understand. However, this is not always true. Miller's approximation removes the feedback promoted by capacitance connecting the input and output nodes. The general consequence of this removal is the unavoidable loss of system's dynamics information. Considering as an example the circuit of Figure 2.26a, the complete analysis shows that the voltage gain is,

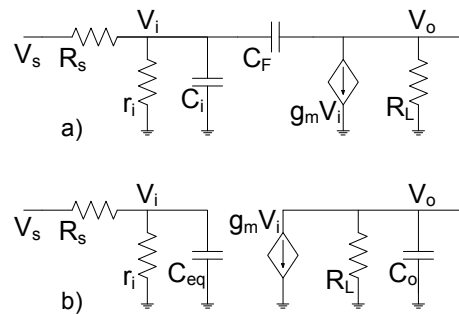


Figure 2.26 - Capacitive feedback; a) complete circuit, b) Miller simplified circuit.

$$\frac{V_o(s)}{V_i(s)} = -g_m R_L \frac{r_i}{r_i + R_s} \frac{1 - s C_F / g_m}{1 + s \left[(1 + g_m R_L) R_{eq} C_F + R_{eq} C_i + R_L C_F \right] + s^2 R_{eq} R_L C_i C_F} \quad (2.60)$$

where $R_{eq} = r_i // R_s$. This voltage transfer functions has one RHP zero and two LHP poles. Applying Miller's theorem to the same circuit results on the simplified version of Figure 2.26b, for which the voltage transfer function is

$$\frac{V_o(s)}{V_i(s)} = -g_m R_L \frac{r_i}{r_i + R_s} \frac{1}{1 + s \left[(1 + g_m R_L) R_{eq} C_F + R_{eq} C_i \right] + s^2 R_{eq} R_L C_i C_F} \quad (2.61)$$

Equation (2.61) reveals the effect of Miller's theorem; the RHP zero has disappeared [111]. In general this zero occurs at very high frequencies and hence, equation (2.61) can be taken as a good approximation of (2.60). However, as seen in the previous section, capacitive feedback can be used to enhance bandwidth. The same sort of bandwidth enhancement is also possible using negative zeros. The essence of this technique is similar to previous discussions, based on the presence of complex poles on the system's transfer function. However, for this case, the effect of the zero has to be combined with resistive feedback. Considering a feedback system with an open-loop transfer function having a RHP zero, it is possible to consider,

$$\begin{aligned} H_{OL}(s) &= H_o \frac{(1 - s\tau_z)}{D(s)} = H_o \frac{(1 + s\tau_z)}{D(s)} \frac{(1 - s\tau_z)}{(1 + s\tau_z)} \\ &\approx \frac{H_o}{D(s)} \frac{(1 - s\tau_z)}{(1 + s\tau_z)} \end{aligned} \quad (2.62)$$

The first transformation in (2.62) is simply a reduction to a minimum phase transfer function, multiplied by a 1st order all-pass transfer function. The second transformation results from the assumption that the zero occurs at higher frequencies than the zeros in $D(s)$. Assuming that $D(s)$ represents a simple pole ($D(s) = 1 + s\tau_i$), and applying negative feedback, the closed-loop transfer function becomes,

$$\begin{aligned} H_{CL}(s) &= \frac{H_{OL}(s)}{1 + \beta H_{OL}(s)} \\ &= \frac{H_o}{1 + \beta H_o} \frac{1 - s\tau_z}{1 + s \frac{\tau_i + \tau_z(1 - \beta H_o)}{1 + \beta H_o} + s^2 \frac{\tau_i \tau_z}{1 + \beta H_o}} \end{aligned} \quad (2.63)$$

Using the maximum flat frequency response conditions for rational transfer functions (given in (2.35)) results in the following constraint on loop gain (βH_o),

$$\beta H_o = \frac{\tau_i^2}{4\tau_z(\tau_i + \tau_z)} \quad (2.64)$$

The bandwidth improvement can be evaluated once the cut-off frequency is known. Then the ratio between cut-off frequencies with and without the RHP zero can be computed. Considering the case with the RHP zero under the maximal flatness constraint, the cut-off frequency results from a 4th order polynomial equation, difficult to express analytically. Alternatively, it is possible to use numerical simulation. Figure 2.27 represents the BWER for this technique using numerical simulation on Matlab. As can be seen the BWER exhibits peaking behaviors similar to capacitive peaking. However for this case the maximum BWER for a fixed zero (see the blue lines) can only be achieved if some frequency response peaking can be tolerated. The maximum flat operation limits the achieved BWER to a nearly 2.15 factor (see the bold red line). As displayed on Figure 2.27 this technique also requires a careful adjustment of the loop-gain (other reported capacitive peaking or inductive peaking techniques do not require this loop-gain adjustment).

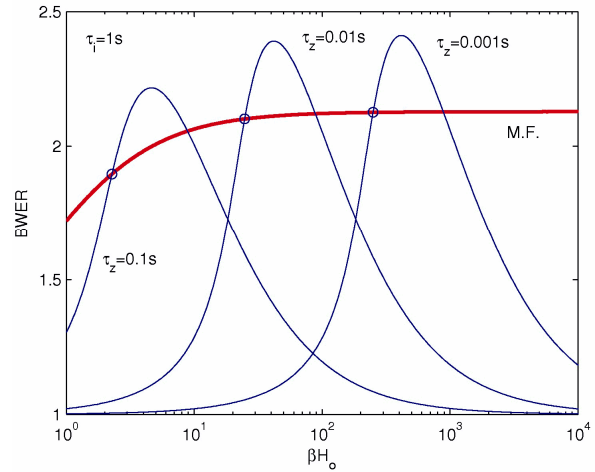


Figure 2.27 - BWER functions, using RHP zeros.

Chapter 3 presents a better exploration of the effect of RHP zeros in feedback amplifiers. As it will be discussed, such RHP zeros are related to delay elements that may occur inside the feedback loop of feedback amplifiers. Traditional design strategies neglect the presence of these delays in feedback amplifiers. Nevertheless, their importance is increasingly relevant as a consequence of the ever demanding gain and bandwidth requirements on actual feedback amplifiers, issues particularly evident in the design of transimpedance amplifiers for optical receiving systems. Chapter 3 presents the concept of delayed feedback, in which the overall design of a feedback amplifier takes into consideration the effect of delay elements around the feedback loop. Frequency response and stability considerations are there discussed, revealing that these delays can be explored as a useful bandwidth enhancement technique in feedback amplifiers of any kind.

2.8 Circuit Level Techniques

On previous sections, several bandwidth enhancement techniques were analyzed, highlighting their major differences and common features. It was shown that maximal flat frequency response

can be achieved for a wide variety of techniques, irrespectively to the initial circuit conception. For instance, a special purpose amplifier can be design using several of the above techniques: it can be a chain of amplifying stages using inductive peaking for each stage, it can be a distributed amplifier, or even a chain of feedback amplifiers. The basic factor allowing these maximal flatness strategies is the presence of complex poles on the overall transfer functions, which can be tuned for optimized bandwidth operation.

Until now circuit level considerations were intentionally left uncovered, because the main effort was to discuss general design strategies suitable for the majority of circuit configurations. However, there are some special circuit configurations able to produce interesting bandwidth enhancement effects. This section presents some of these circuits that due to their widespread usage deserve a brief presentation.

2.8.1 Cascode Amplifier

There are several transistorized amplifying stages suitable for high bandwidth operation [247, 252, 264, 276, 277, 279, 284]. Between the three basic amplifying stages, CX (common X, also known as common emitter or common source for bipolar and FET technologies respectively), CY (common Y, also known as common base or common gate) and CZ (common Z, also known as common collector or common drain), CX and CY are able to produce voltage gains larger than unity. It is also well known that CY configurations have larger bandwidth and smaller input impedances than their CX counterparts. A better amplifying configuration joining the benefits of both CX and CY configurations (large bandwidth, large gain and large input impedance) is the Cascode amplifier, composed by a CX input stage followed by a CY output stage [122, 123].

Figure 2.28 shows the CX amplifying stage and its small signal equivalent. Using Miller's approach¹⁰, the capacitance c_z can be converted into two capacitances, one placed at the input in parallel with c_x and other at the output, with values approximately given by $(1+g_m R_L)c_z$ and c_z respectively. The voltage gain is given by,

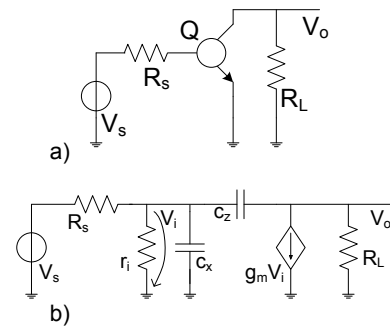


Figure 2.28 – Common X configuration.

¹⁰ For this case it is possible to simplify circuit analysis using Miller's approach, since the effect of the RHP zero due to c_z is not taken into consideration.

$$\frac{V_o(s)}{V_s(s)} \approx -\frac{g_m r_i R_L}{r_i + R_s} \frac{1}{(1 + s R_{eq} C_i)(1 + s R_L C_z)} \quad (2.65)$$

where $R_{eq} = r_i // R_s$ and $C_i = c_x + (1 + g_m R_L) c_z$. Equation (2.65) shows two poles, one due to the input circuit and other due to the output. In general the output pole occurs for very large frequencies, thus the bandwidth of the CX stage is dominated by $R_{eq} C_i$. The Miller's multiplying effect reduces the bandwidth since C_i increases with the voltage gain $g_m R_L$. An effective bandwidth enhancement technique is to reduce the voltage gain of the stage.

However, a more expedite technique is to use a cascade association of one CX stage with one CY stage, forming the well known Cascode stage. Figure 2.29 represents the Cascode amplifier together with its small signal equivalent. This representation assumes equal transistors, and thus equal small signal equivalent circuits. A straight forward analysis shows that the overall

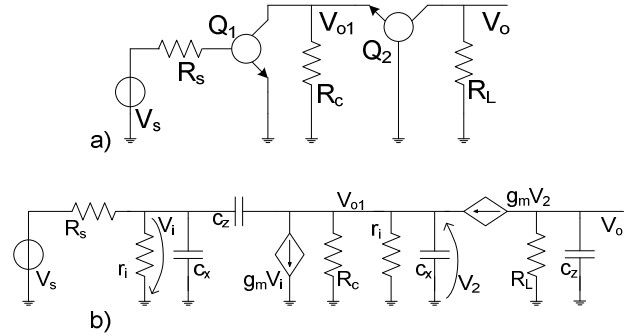


Figure 2.29 - Cascode amplifying stage.

voltage gain transfer function can be divided into three transfer functions due to the input circuit, to the middle stage and to the output stage. Thus,

$$\begin{aligned} \frac{V_o(s)}{V_s(s)} &= \frac{V_i(s)}{V_s(s)} \frac{V_{o1}(s)}{V_i(s)} \frac{V_o(s)}{V_{o1}(s)} \\ &\approx \left(\frac{r_i}{r_i + R_s} \frac{1}{1 + s R_{eq1} C_i} \right) \left(-\frac{g_m (r_i // R_c)}{1 + g_m (r_i // R_c)} \frac{1}{1 + s R_{eq2} C_m} \right) \left(\frac{g_m R_L}{1 + s R_L C_z} \right) \end{aligned} \quad (2.66)$$

where $R_{eq1} = r_i // R_s$. The benefits of this association are revealed firstly by realizing that R_{eq2} (the effective load of the first stage) is approximately $1/g_m$. The impact of this resistance on both Miller equivalent capacitances C_i and C_m is,

$$\begin{aligned} C_i &= c_x + (1 + g_m R_{eq2}) c_z \\ &\approx c_x + 2c_z \\ C_m &= c_x + \left(\frac{1 + g_m R_{eq2}}{g_m R_{eq2}} \right) c_z \\ &\approx c_x + 2c_z \end{aligned} \quad (2.67)$$

Equation (2.66) reveals that the DC voltage gain is similar to the DC voltage gain of a simple CX stage (since in general $g_m (r_i // R_c) \gg 1$), however, the input capacitance C_i is essentially fixed thus assuring higher bandwidths. Assuming that the cut-off frequency of both stages is essentially set by the input circuitry, the BWER benefit in using Cascode association instead of simple CX stages is

given by,

$$BWER = \frac{(R_{eq}C_i)_{CX}}{(R_{eq1}C_i)_{CX-CY}} = \frac{1 + (1 + g_m R_L) \frac{C_z}{C_x}}{1 + 2 \frac{C_z}{C_x}} \quad (2.68)$$

Equation (2.68) shows that for high voltage gain requirements the Cascode circuit is indeed more bandwidth effective than a simple CX stage, since the BWER increases with the voltage gain factor $g_m R_L$. In this case both gain and bandwidth are enhanced by trading the simplicity of the CX stage by a more evolved Cascode stage (CX-CY).

2.8.2 Cherry-Hooper Amplifier

When both high voltage gain and high bandwidth are required, the best choice to accommodate both is to divide the gain by an appropriate number of cascaded stages with sufficiently high bandwidth. The simplest possible configuration uses standard CX amplifiers as the basic amplifying stage. As discussed on the previous section, this strategy is limited by Miller effects. A better possibility is to use Cascode stages instead of the simple CX stages: the resulting gain is nearly the same on a per-stage basis, but the bandwidth is larger due to the Miller effect reduction property of this configuration. Other possibility is to use feedback. The Miller effect reduction on Cascode amplifiers is achieved through a gain reduction of the first CX stage; this is accomplished using a second CY stage with low input impedance.

The same effect can be achieved using a transadmittance-transimpedance (commonly referred as TAS-TIS). The first stage can use a CX stage to provide transadmittance gain, and the second stage can use a CX stage with shunt-shunt feedback to provide the transimpedance gain (see Figure 2.30c) – the low input impedance of the shunt-shunt CX stage effectively reduces the Miller effect of the first stage. This is the basic constitution of the Cherry-Hooper amplifier [117, 279, 284].

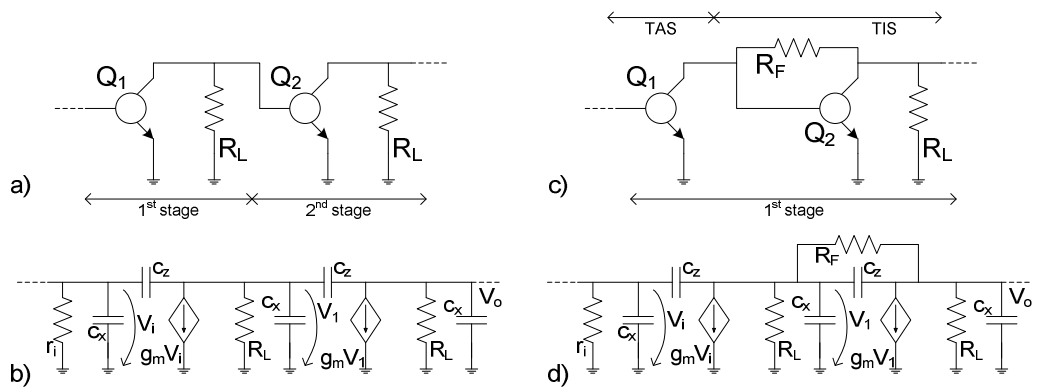


Figure 2.30 - Cherry-Hooper amplifier: a) and b) traditional CX chain amplifier (small signal and detailed small signal equivalents respectively); c) and d) Cherry-Hooper TAS-TIS stage (same).

Figure 2.30a and c, shows the reference amplifier composed by a chain of CX stages, and the alternative solution of the Cherry-Hopper amplifier; for both cases it is also represented (Figure 2.30b and d) the small signal equivalent circuit assuming equal transistors per stage. In order to compare the enhanced performance of the Cherry-Hopper configuration it is only necessary to consider the transfer function between V_i and V_o for a single stage. This means that the effect of the intrinsic capacitances c_x and c_y due to the input transistors can be ignored. Considering first the CX chain, the transfer function of a single stage is given by,

$$\frac{V_o(s)}{V_i(s)} \approx -\frac{g_m R_L}{1 + s R_L C_i} \quad (2.69)$$

where R_L may include the effect of both transistors input and output resistances and C_i is given by $C_i = c_x + (1 + g_m R_L) c_z$ as above. Considering now the Cherry-Hopper stage (Figure 2.30d), the transfer function of one stage is given by,

$$\frac{V_o(s)}{V_i(s)} = \frac{A_o (1 + b_1 s)}{1 + a_1 s + a_2 s^2} \quad (2.70)$$

where,

$$\begin{aligned} A_o &\approx g_m R_F \\ b_1 &= -\frac{c_z}{g_m} \\ a_1 &\approx \left(2 + \frac{R_F}{R_L}\right) \frac{c_x}{g_m} + R_F c_z \\ a_2 &\approx \frac{R_F}{g_m} \left(1 + \frac{c_z}{c_x}\right) c_x^2 \end{aligned} \quad (2.71)$$

The approximated results of equation (2.71) are valid under the assumption that $g_m R_F \gg 1$ and $g_m R_L \gg 1$. Neglecting the RHP zero on (2.70), maximum flat frequency response is obtained if $a_1^2 = 2a_2$. Under this assumption and taking the same DC gain for both configurations ($R_F = R_L$), the cut-off frequency is determined by a_2 alone. The resulting BWER of the Cherry-Hopper configuration becomes,

$$BWER = \frac{(f_c)_{CH}}{(f_c)_{CX}} = \left(1 + (1 + A_o) \frac{c_z}{c_x}\right) \sqrt{\frac{A_o}{1 + c_z/c_x}} \quad (2.72)$$

As equation (2.72) shows the bandwidth benefit of using Cherry-Hopper amplifying stages instead of simpler CX stages, is at least proportional to the square root of the desired gain for transistors with low c_z , c_x ratios. There is however a restriction on the maximum value of A_o ; since the design uses a maximum flat design approach, A_o becomes restricted by the ratio between c_z and c_x .

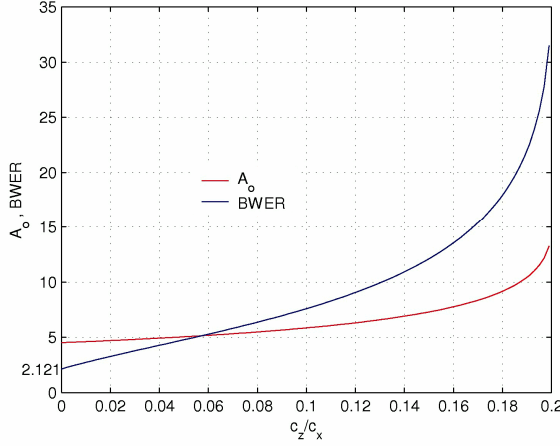


Figure 2.31 - BWER and gain of the of the Cherry-Hooper amplifier.

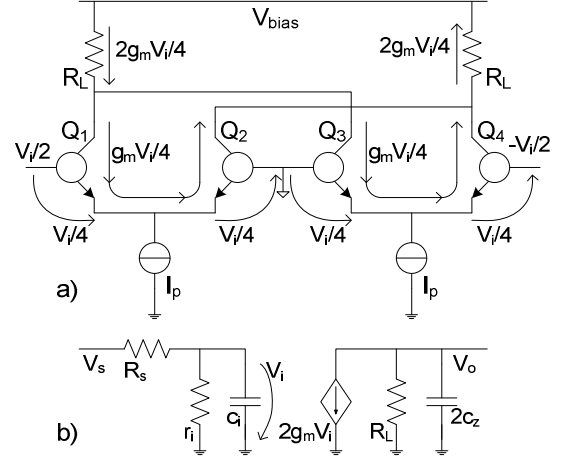


Figure 2.32 - f_T Doubler: a) functional concept; b) differential mode half circuit.

Figure 2.31 depicts both gain (A_o) and BWER of the Cherry-Hooper against the CX amplifiers. The ratio of c_z and c_x is also restricted for flat operation (c_z/c_x larger than 0.2 prevents maximum flatness). For values of c_z/c_x inside the range between 0 and 0.2 the BWER is always larger than 2 and the gain per-stage is larger than 4.5.

Larger gain values can be achieved using improved versions of the basic Cherry-Hooper amplifier. A simple modification is to introduce an output buffer, allowing larger values of R_L . However, this technique has the drawback of enhancing the capacitive load of the transimpedance stage [124-131]. The Cherry-Hooper amplifier is often used in differential amplifiers as a means of extending the bandwidth; several examples of this technique can be found in [124-131].

2.8.3 f_T Doubler

Certain transistor configurations exhibit an interesting property of doubled f_T . One such case is the circuit of Figure 2.32, known as f_T doubler [279, 284]. This circuit consists of a differential structured constructed around two differential pairs. The inputs of the differential pairs are connected in series, while the outputs are cross-connected in parallel. Assuming differential signal operation, the middle point of the differential pairs is at a fixed potential imposed by the bias current sources. For equal transistors, the input signal divides itself by the four YX input ports of the transistors. Since the current entering the Z terminal of one transistor is controlled by the signal voltage between YX terminals (true for both FET and Bipolar transistors), the current contribution of each transistor is exactly $g_m v_i/4$. Due to the parallel cross-connection, transistors Q_1 and Q_3 contribute to a total current of $2g_m v_i/4$, while transistors Q_2 and Q_4 contribute to a total current of $-2g_m v_i/4$. The net result is the same output differential current as a simple differential pair. However, an analysis of the small signal equivalent half circuit, shows that the transconductance of transistors

Q_1 and Q_3 (Q_2 and Q_4 alternatively) add together, resulting in a small signal equivalent circuit resembling a single transistor (as for a simple differential pair) with doubled g_m (see Figure 2.32). Evaluating the correspondent transition frequency for this situation, results in a doubled f_T . The same f_T doubling property is also present in Darlington configurations [279, 284].

Doubled f_T does not mean necessarily larger bandwidth. For this particular case, having a doubled g_m also implies nearly doubled Miller effect on c_z , which corresponds to smaller cut-off frequencies. However, superior behavior results if frequency peaking techniques are employed together with this circuit.

2.8.4 *Current-Mode Strategies*

A wide variety of amplifier design strategies are based on two important concepts: feedback and the nullor.

The nullor introduced by Carlin and Youla in 1961 [116, 118] and explored by Tellegen [119], became widely accepted as the basic ideal element in the synthesis of feedback amplifiers. The nullor is a two-port network consisting of a nullator at its input and a norator at its output. The nullator imposes a virtual short circuit at the input circuit of the nullor; that is, the input voltage is zero as in a short circuit, but no current flows. The norator output behaves in opposite fashion, leaving both output voltage and current unrestricted. It soon became apparent that this network element was the adequate personification of the voltage-mode operational amplifier (VOA), named several years later by Raggazzini [115]. However, the first integrated VOA appeared only in 1963; the $\mu A702$ invented by Widlar, soon followed by the $\mu A709$ [120, 121]. This new element was capable of synthesizing the nullor behavior through a high voltage gain between the output and the difference between two inputs. The relation of voltages clearly justifies its voltage-mode character. The rapid development of the VOA pointed researcher's interests towards voltage-mode solutions. The first true current-mode nullor was the second generation current conveyor (CC-II), which appeared only in 1970, issued by Sedra and Smith (after the first generation current conveyor developed in 1968 by the same authors [215, 216]). By that time, voltage-mode philosophy was so deeply embedded in designs that this new element did not deserve much attention. Only several years later, with new trends of low-voltage operation, had these current-mode devices occupied a well deserved recognition [259, 276].

Current-mode design is known to produce circuits with large bandwidths. The philosophy behind current-mode design is to treat signals as currents instead of voltages. This philosophy is well suited for integrated circuit design, where the medium is pre-dominantly capacitive: several metal layers act together with isolation materials as distributed parasitic capacitances all over the circuit.

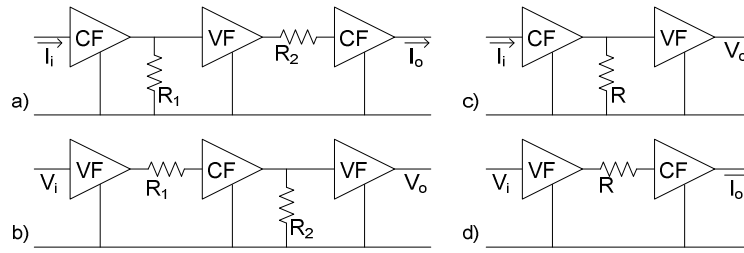


Figure 2.33 - Current-mode amplifying topologies: a) current amplifier; b) voltage amplifier; c) transimpedance amplifier; d) transadmittance amplifier.

Physical restrictions arise due to the fact that a signal voltage has to waste time charging or discharging these parasitic capacitances, and thus large voltage swings also have large amounts of time wasted. If instead of voltages, signals are treated as currents flowing on a circuit with low impedance nodes, this problem is effectively minimized. The node voltages will, for the majority of applications, exhibit low swings and the charging-discharging times are thus reduced, allowing higher bandwidth operation.

Figure 2.33 presents the current-mode perspective of the four amplifier topologies, current, voltage, transimpedance and transadmittance (from a) to d) respectively). Assuming that high bandwidth voltage and current followers (VF and CF, for short) are available, signal amplification is accomplished using voltage to current or current to voltage conversions on simple resistances. Current amplification is accomplished using a current to voltage conversion in a resistance followed by voltage to current conversion by another resistance, as depicted on Figure 2.33a. The final current gain is given directly by the ratio of the two resistances (R_1/R_2 for the case of Figure 2.33a). Voltage amplification is accomplished in a similar manner, taking first a voltage to current conversion followed by a current to voltage conversion. For the case of Figure 2.33b) the final gain is directly proportional to R_2/R_1 . Transimpedance (transadmittance) gain needs one single current to voltage (voltage to current) conversion (Figure 2.33c and d). In all the four cases of Figure 2.33, the achieved bandwidth is entirely dependent on the bandwidth of both the voltage and current followers assuming ideal input/output impedances.

This simple explanation shows how current-mode amplifiers can achieve constant bandwidth operation independent of the gain level. This same property is not observed in voltage-mode amplifiers: using feedback configurations and VOAs, the GBW of the closed-loop amplifier becomes restricted, implying that large bandwidths are achieved through severe gain penalties. The real picture for current-mode amplifiers is not as perfect as the above description. The presence of non-ideal input/output impedances on both voltage and current followers makes the current or voltage conversion frequency dependent. In general, the bandwidth is limited by these parasitic impedances together with load resistances used for conversion purposes (assuming that both current

and voltage followers have large bandwidth capabilities).

Chapter 5 explores the concept of current matching as a means of bandwidth enhancement in transimpedance amplifiers. This concept is based on the conception of adequate current-matching-devices (CMD) able to adapt the input of a transimpedance amplifier to a capacitive current source, often found in optical receiving systems. Such CMDs are characterized by an unusually small input impedance (thus reducing the bandwidth penalty posed by the capacitive source), and an idealized current output. Current-mode design strategies employing current-conveyors are suggested as the most appropriate to the conception of these CMDs.

2.9 Concluding Remarks

This chapter discussed the general GBW limitations on amplifier design. The theoretical foundations for GBW limits in amplifiers were originally developed by Bode. According to Bode's results, the ratio g_m/C specifies the GBW limit of one transistor. Since amplifiers are composed of several transistor associations, this is also a quantity of interest for amplifier design.

Design methods able to approach or even surpass these limitations were also presented and discussed. This chapter adopted a classification based on the general characteristics of these methods. According to this classification, there can be conceptual and circuit level techniques, able to address gain, bandwidth or both gain and bandwidth problems in amplifiers. The description of these methods adopted a simplified approach, revealing the philosophy behind each method. More information on each of the discussed methods can be found on the references furnished with this thesis.

3 DELAYED FEEDBACK AMPLIFIERS - THEORY

“Hey Pal! How do I get to town from here? And he said: Well just take a right where they're going to build that new shopping mall, go straight past where they're going to put in the freeway, take a left at what's going to be the new sports center, and keep going until you hit the place where they're thinking of building that drive-in bank. You can't miss it. And I said: This must be the place.” - Laurie Anderson.

Summary

This chapter presents the concept of delayed feedback amplifiers as a new method of improving bandwidth in feedback amplifiers. A discussion on the benefits and drawbacks of this method is addressed taking as reference other bandwidth enhancement techniques presented on chapter 2. Techniques suitable for the design of maximal flat frequency response amplifiers employing delayed feedback are also presented, leading to the principal result of the chapter. Finally, stability and robustness issues are discussed.

3.1 Delayed Feedback Amplifiers

Several bandwidth enhancement techniques suitable for high gain and bandwidth amplifiers were presented and discussed on chapter 2. It was shown that the great majority of these techniques are based on the optimal placing of the amplifier's poles. In general the amplifier has a rational transfer function consisting of a ratio of two polynomials. A key bandwidth optimization procedure consists in shaping these polynomials in order to achieve maximum flat frequency response. It was shown that this shaping resort to pole-zero placement procedures intimately related with circuit's time constants. Depending on the degree of these polynomials these might imply a tight control of several time constants. Since the circuit's time constants are directly related with one or several bandwidth enhancement elements (see for instance the case of inductive peaking, on section 2.6.2), this tight control can be difficult as the circuit's complexity increases. As a result, instead of the desired flat frequency response bandwidth enhancement, frequency peaking can often occur.

Implementation details further add design complexity to these frequency shaping procedures, in particular when the frequency optimization is retrieved by resorting to the usage of inductances. Integrated circuits environment is predominantly capacitive, thus meaning that the design of any inductance is also affected by interlayer parasitic capacitances. Adding one or more inductances to the amplifier's circuit must be accomplished using complex and time consuming simulation steps, since in general each added inductance has its own transfer function dependent on several process parameters. The impact on the final circuit is at least unpredictable within a certain confidence interval. In general one of two phenomena may result: slightly less bandwidth than predicted or, unwanted frequency peaking. Compensation of these behaviors is difficult to accomplish without the usage of complex control circuitry, which in general result in further bandwidth reduction.

Negative Feedback techniques are a natural answer to these problems. One important feature of negative feedback amplifiers is their inherent ability to compensate external perturbations. By external perturbations it is meant all the factors that can affect the desired amplifier response in a non-productive manner: temperature and process variations are examples of such external perturbations. On the particular case of amplifier's design, these external perturbations can exhibit different manifestations. Gain loss, bandwidth loss or frequency peaking are without doubt the most challenging effects affecting the gain-bandwidth optimization problem. Negative feedback amplifiers can minimize the impact of these effects due to the feedback action; an increase in gain is automatically cancelled by an increased feedback signal, thus contradicting the original cause. Since in feedback amplifiers gain and bandwidth are closely related, gain compensation also results in bandwidth stabilization. This line of reasoning clearly justifies the adequacy of feedback amplifiers

for situations where stable and immune operation is demanded.

Another important feature of negative feedback amplifiers is their inherent ability to supply a simple and unique form to implement compensation mechanisms. If the frequency response does not display the desired characteristics, it is always possible to adjust slightly the gain and compensate the deviations. The gain control mechanism is the true essence of feedback amplifiers: the closed-loop gain is controlled by the feedback loop, it is only necessary to add some control in the feedback path.

Another possibility is to explore the presence of delays in feedback amplifiers as a means to achieve bandwidth improvements. Recently published contributions justify this possibility. On [91, 107, 108, 112] the usage of distributed elements on the feedback loop of these amplifiers lead to significant bandwidth improvements. The presence of RHP zeros inside the feedback loop of these amplifiers can also be explored as a bandwidth enhancement technique [298, 299]. These examples suggest the presence of delay elements associated with the feedback path. In fact distributed resistances are a simple approximation of a lossy transmission line and its associated time delay [107]. It is also possible to show that RHP zeros can be viewed as delay elements [248, 268].

Nevertheless, the effect of delays in amplifier circuits as detrimental elements is well established since the early days of electronics [77, 83, 102, 195, 214, 248]. Delays can be regarded as excess phase terms; their recognized effects on electronic circuits are mostly: phase distortion – signal components of different frequencies are affected with different delays; and instability – in feedback circuits, the presence of an excess phase generator inside the feedback loop can act as an destabilizing element, increasing the phase argument towards the critical 180° . Design procedures able to compensate the negative effects of these delay elements in valve amplifiers have been reported during 1940 by Shaw and Bode [77, 248]. Valve amplifiers were particularly affected by the presence of delays due to the physical dimensions of the circuits, since the electrical paths on the circuit were long enough to produce delays comparable to the circuit's time constants [77]. The first transistorized circuits also considered delays due to the transistor operation: the early bipolar transistor models usually included a complex exponential term associated with the alpha current gain [83, 195, 214]. However, due to the rapid evolution of transistors and integrated circuits, the effect of small delays became negligible. In fact circuit dimensions became much smaller. The effect of delays became appreciable only in very few and particular situations; high gain and bandwidth transimpedance amplifiers are one such case. Recently, Nowack [102] discussed delay induced instability problems in transimpedance amplifiers used for high speed fiber optic communications systems. The combination of high gain and high bandwidth requirements in feedback transimpedance amplifiers is by itself a difficult one from the stability point of view.

Stability problems become even worse if an excess phase term is also present.

Few works explored in an explicit fashion the presence of delays as beneficial elements in electronic circuits [80, 86]. It is not surprising to notice that the effect of delays in feedback amplifiers is not very different from the effect of other frequency peaking techniques. In fact, a delay element in the feedback loop of a feedback amplifier can be used to move adequately the closed-loop dominant poles to some prescribed position. The major difference lies on the fact that the dynamics of delayed feedback amplifiers¹¹ contain an infinite number of poles, rendering the design procedure extremely complicated. Explorations of this simple observation were originally reported by Nero *et al* in [296, 297]. Concerning the design of delayed feedback amplifiers two problems arise with significant importance: the maximal flat design problem and the stability problem. Both of them are complex problems. Delayed feedback amplifiers are ruled by characteristic equations containing both polynomials and complex exponential terms. This type of characteristic equations fall inside a large category of equations denominated quasi-polynomial equations. As stated on chapter 2, maximal flat design procedure for amplifiers having polynomial dynamics is a simple matter; it suffices to apply recursively theorem 2.1 to the amplifier's transfer function. Since in this case the transfer function is composed by polynomials of known order this correspond to a finite set of design equations. The same is not true for quasi-polynomial dynamics, because complex exponentials have an infinite number of non-zero derivatives. An efficient approach able to simplify this problem is later advanced in this chapter.

To study the stability restrictions in delayed feedback amplifiers it is also possible to use some general results of modern control theory concerning time delay systems. The topic of feedback systems with time delays is a very old one [282]. The first mention to time delay systems appeared in the eighteen century, associated to functional differential equations (FDE). Several examples were studied by mathematicians as Euler, Bernoulli, Lagrange, Laplace and Poisson. In the early twentieth century, several problems were again modeled using FDEs. Examples of these modeling approaches can be found in various areas of science: viscoelasticity problems, predator-prey models in population dynamics and ship stabilization, to mention some. The emergence of modern control theory as a discipline brought several new perspectives on this kind of systems. The problem of stability in time delay systems become an important one. In 1942, Pontryagin presented the first fundamental results on the zeros of quasi-polynomials often found in time delay systems [132, 249]. This opened the way for all the future advances concerning stability in this kind of systems. Several

¹¹ From now on the term delayed feedback amplifier is used to mention a feedback amplifier comprising one explicit delay element inside its feedback loop.

techniques were advanced concerning time domain, frequency domain and state-space methods. The frequency domain approach is from the amplifier's design perspective the most direct related technique to this work. Among the frequency domain techniques, several results can be applied to the stability assessment in time delay systems. Some examples of such techniques are: the Nyquist stability criterion [257, 283]; analytical root-locus techniques [133-135], generalizations of the classical Hermite-Biehler theorem to quasi-polynomials [132, 149-155, 285]¹², and generalizations of Bode's integral relations [136, 137], among others [282].

This chapter presents a unified theory suitable to bandwidth optimization of feedback amplifiers employing phase shaping elements (both delay and all pass transfer functions, or even both), in their feedback loop. The essence of this theory is based on the bounded behavior of the closed-loop frequency response of feedback amplifiers. An exploration of this bounded phenomenon not only reveals the maximum achievable bandwidth for a specified feedback amplifier, but also exposes an efficient method of attaining this maximum.

3.1.1 General Discussion

Figure 3.1 represents a general model of a feedback amplifier, comprising a forward amplifier with transfer function $A(s)$ and a feedback path with transfer function $\beta(s)$. The feedback type is

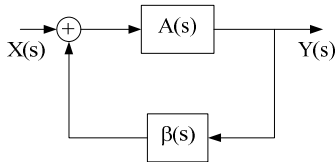


Figure 3.1 - Reference feedback amplifier.

defined from the combination of the input signal $X(s)$ with the feedback signal, $\beta(s)Y(s)$. If the combination is additive, the feedback is positive or regenerative, and the feedback signal reinforces the input signal. If on the other hand, the feedback signal $\beta(s)Y(s)$, subtracts itself to the input signal, then the feedback is negative or degenerative, and the

feedback contradicts the original cause. Both feedback types can be used in amplifier design: negative feedback is often used to produce stable and robust amplifiers; positive feedback is used in order to enhance bandwidth, or even to design oscillators. Depending on the phase variation of the loop-gain, $G(s)=\beta(s)A(s)$, both feedback types can occur within the same amplifier at different frequencies. Assuming negative feedback, the closed-loop transfer function is for the amplifier in Figure 3.1 given by,

¹² The classical Hermite-Biehler theorem states the necessary and sufficient conditions for polynomials to be stable in Hurwitz sense. A direct manipulation of the Hermite-Biehler theorem leads to the well known Routh-Hurwitz stability criterion [283].

$$H(s) = \frac{Y(s)}{X(s)} = \frac{A(s)}{1+G(s)} \quad (3.1)$$

The correspondent frequency response can be obtained restricting the complex variable s to the imaginary axis. Considering the squared magnitude of the frequency response,

$$|H(j\omega)|^2 = \frac{|A(j\omega)|^2}{|1+G(j\omega)|^2} \quad (3.2)$$

It is a simple matter to show that (3.2) is a bounded function of ω , resulting in,

$$|H_{Lo}(j\omega)|^2 = \frac{|A(j\omega)|^2}{(1+|G(j\omega)|)^2} \leq |H(j\omega)|^2 \leq \frac{|A(j\omega)|^2}{(1-|G(j\omega)|)^2} = |H_{Up}(j\omega)|^2 \quad (3.3)$$

This simple observation reveals much about the bandwidth enhancement possibilities for feedback amplifiers having a prescribed set of open-loop poles and zeros. Assuming that it is possible to improve the closed-loop bandwidth for a desired closed-loop gain without interfering with open-loop gain and time constants, then (3.3) states that the maximum bandwidth improvement which can be achieved is limited by an upper bound magnitude function. This argument excludes the capacitive peaking techniques described in chapter 2, since they imply reshaping the open-loop poles. The only possibility is to consider the presence of phase shaping elements, as is for instance the case of non-minimum open-loop transfer functions, or open-loop transfer function containing delays. The first case is present when the open-loop transfer function has RHP zeros; this can be used to improve bandwidth as discussed on chapter 2. A second, more general example has open-loop transfer function containing delay elements. A general observation

shows that, the presence of non-minimum phase terms or explicit delay elements does not affect the magnitude of the open-loop transfer function of a feedback amplifier. The following discussion will ascertain that these two examples are indeed explained efficiently by the same framework.

Consider the simple example of a feedback amplifier having open-loop gain $A_o=1000$, with only one open-loop pole at 1 rad/s , and a constant feedback factor $\beta=0.1$. Figure 3.2 represents the closed-loop

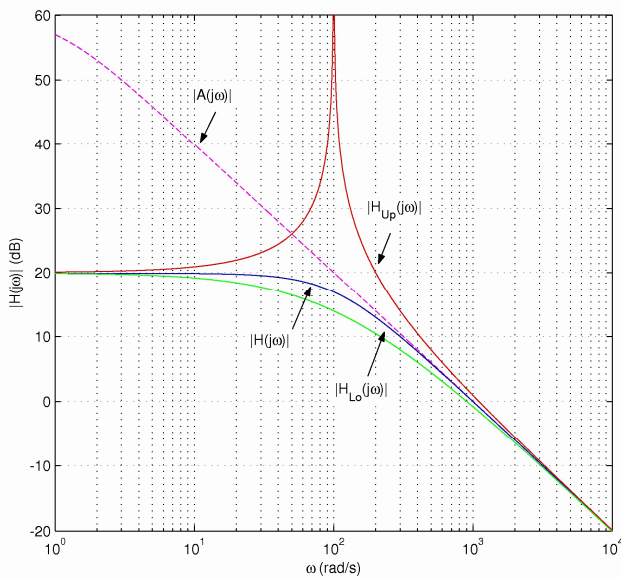


Figure 3.2 - First order feedback amplifier.

response for this example. The two associated bounding functions of (3.3) are also present, $H_{Up}(\omega)$ and $H_{Lo}(\omega)$, respectively the upper and lower bounds. The closed-loop frequency response follows the open-loop high frequency asymptote. The upper bounding function states that it is possible to approach this same asymptote from higher gain values. This situation would correspond to an enhanced bandwidth version of the same amplifier.

A possible method to approach the upper bound cut-off frequency is to introduce phase shaping elements inside the feedback loop. Using the above considerations, this can be accomplished using open-loop transfer functions containing delays or all-pass terms. Introducing a pure delay element on the feedback loop, results on an open-loop transfer function given by,

$$G(s) = \beta(s) A(s) e^{-sL} \quad (3.4)$$

where L is positive number representing an effective time delay of L seconds. On the other hand, all-pass terms can be used to simulate the effect of a pure delay element. Indeed all-pass transfer functions can be viewed as Padé approximants of the complex exponential [268]. Assuming for simplicity a first order all-pass approximation of the delay, the open-loop transfer function becomes,

$$G(s) = \beta(s) A(s) \frac{1-sT}{1+sT} \quad (3.5)$$

where T is a positive number representing an approximated delay of $2T$ seconds. As can be seen in (3.5), the open-loop transfer function comprises a RHP zero and LHP pole of equal magnitude. It is now clear that, both transfer functions in (3.4) and (3.5) have the same magnitude with and without the delay elements.

Figure 3.3 represents the magnitude of the closed-loop frequency response of the first order

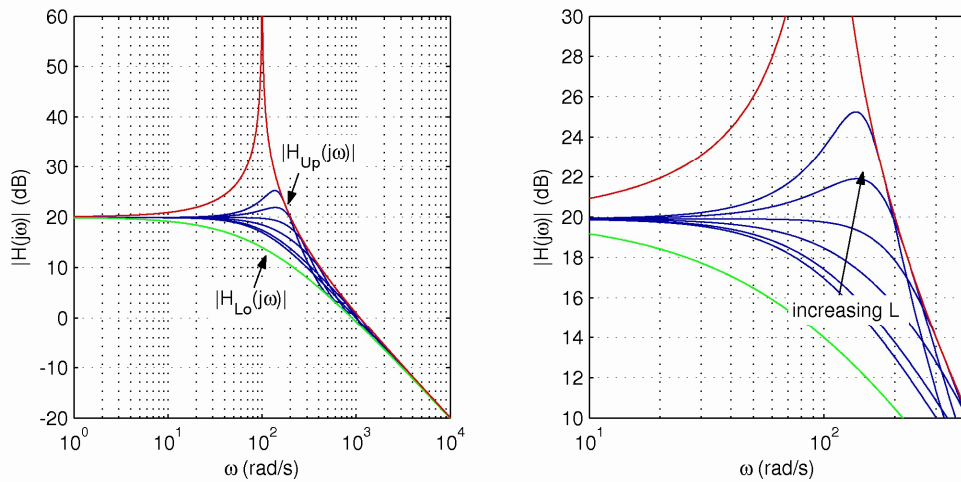


Figure 3.3 – Closed-loop magnitude response adding a delay element to the feedback loop (plus detailed zoom of the cut-off region).

system of the previous example adding a delay to its feedback loop, as in (3.4). As can be seen the cut-off frequency is strongly dependent on the delay value; increasing the delay within certain bounds allows significant bandwidth improvements. However, the delay cannot be increased without bound since this would result in two consequences: first, frequency peaking phenomena appear (this is also present on Figure 3.3); and then loss of stability. Nevertheless, it is possible to tune the delay for maximal flat frequency response operation and at the same time, retrieve some bandwidth enhancement, as shown on Figure 3.3.

The cut-off frequency has to be computed in order to assess the maximum bandwidth improvement. This is a slightly complex procedure, since the closed-loop dynamics include a delay element. Considering the same example as above, the cut-off frequency may be determined using the standard 3dB crossing procedure, or equivalently,

$$|H(j\omega)|^2 = \frac{|H(0)|^2}{2} \quad (3.6)$$

Using the above parameters, equation (3.6) becomes,

$$\left[\frac{1+G \cos(\omega L)}{1+G} \right]^2 + \left[\frac{a_1 \omega - G \sin(\omega L)}{1+G} \right]^2 = 2 \quad (3.7)$$

Where G represents the loop-gain given by $G=\beta A_o$ (assuming β fixed), $1/a_1$ represents the pole of the open-loop transfer function, and A_o its small frequency gain. Equation (3.7), shows that the cut-off frequency is now the solution of a non-linear equation in ω . The closed form solution of (3.7) is not known, but the values of the cut-off frequency can be computed using numerical procedures. Since the variable of interest is the bandwidth improvement against the standard non delayed feedback amplifier, some useful transformations can be applied to (3.7). The first transformation consist in the simple variable change $\alpha=\omega L$, where α is now the normalized frequency variable measured in radians. A second transformation allows the direct evaluation of the bandwidth improvement; as in chapter 2 the bandwidth improvement can be conveniently expressed by the BWER defined by the ratio between the cut-off frequencies of the system under analysis against the reference system. Considering the reference system, the non-delayed version of the feedback amplifier, with cut-off frequency expressed by $\omega_r=(1+G)/a_1$. Using all these considerations, (3.7) becomes,

$$\frac{1+G^2}{(1+G)^2} + \frac{2G}{1+G} \left[\frac{\cos(\Phi)}{1+G} - K \sin(\Phi) \right] + K^2 = 2 \quad (3.8)$$

where K represents the BWER and Φ is a new variable defined as $\Phi=K(1+G)\delta$, with δ being the pole-delay product, $\delta=L/a_1$.

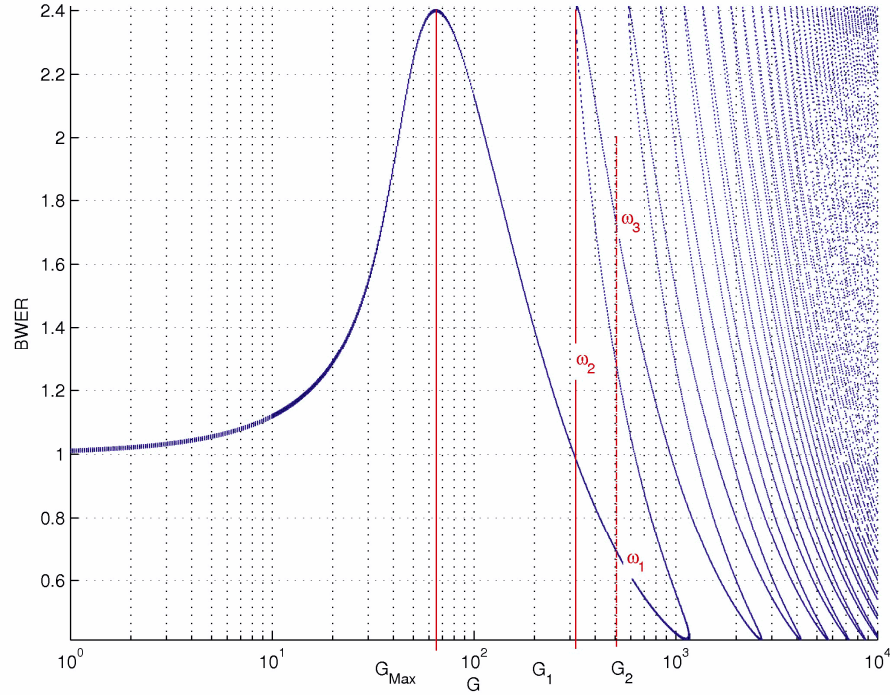


Figure 3.4 – BWER as function of the loop-gain for a fixed pole-delay product ($\delta=0,01$).

Two procedures can be applied to compute the BWER: fixing the loop-gain and sweeping the pole-delay product, or, alternatively fixing the pole-delay product and sweeping the loop-gain. As it will be shown later (on section 3.5), a suitable procedure for the design of delayed feedback amplifiers comprises a graphical representation of the BWER against the pole-delay product and one equation to compute the loop-gain. For now, it is instructive to inspect the BWER function for fixed pole-delay products as the loop-gain is swept. Since (3.8) comprises trigonometric sine and cosine functions it seems logical to expect multiple solutions for a single (G, δ) pair. This observation turns the numerical procedure difficult to execute. An alternative is to use (3.3) to restrict the values of BWER for each (G, δ) pair, and (3.8) as checking criterion. If the set of BWER values has a large number of points, the resulting BWER curve will give an accurate solution of (3.8).

Figure 3.4 depicts the BWER as a function of the loop-gain with a fixed pole-delay product of $0,01$ (assuming an open-loop pole of 1 rad/s). As predicted, the BWER is a multi-valued function for some (G, δ) conditions. This multiple value behavior arises due to the fact that the frequency response may exhibit multiple maxima, thus meaning the occurrence of multiple 3dB crossings. As Figure 3.4 suggest there are two regions of interest, the first being the single cut-off region established for loop-gain values smaller than G_1 . The second region represents the multiple cut-off behavior, established for loop-gain values larger than G_1 . The maximum BWER attains a value near 2.4 when the loop-gain is G_{max} and tends to a slightly larger value for larger values of the loop-gain

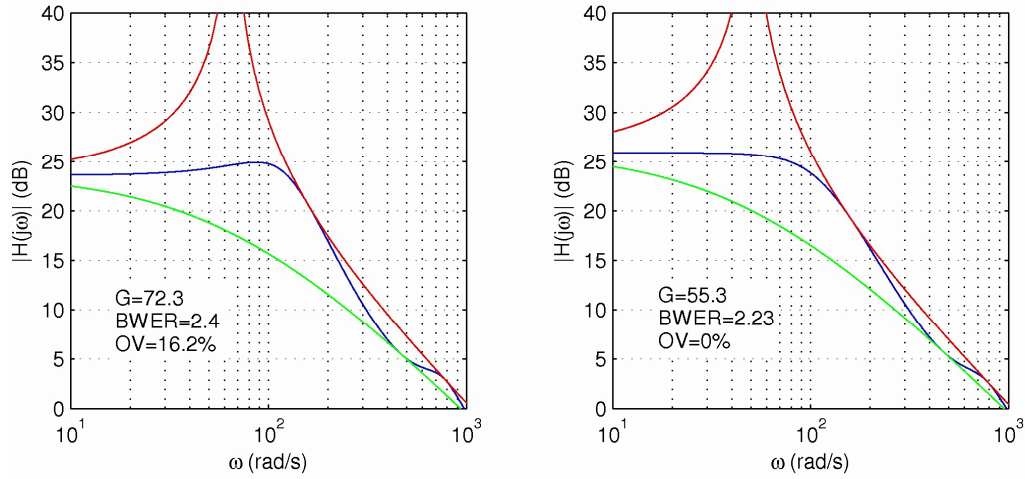


Figure 3.5 - Closed-loop frequency response comparison for maximum BWER and maximum flat conditions.

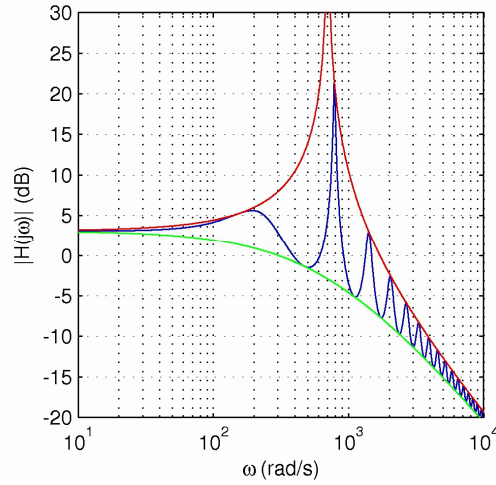


Figure 3.6 – Closed-loop frequency response with multiple peaks.

(it can be shown theoretically that the maximum achievable BWER is $1+2^{1/2}$, see appendix B.1). For values of the loop-gain beyond G_{max} , the amplifier's stability becomes seriously compromised as the peaking becomes more and more pronounced. In fact, even for the maximum BWER condition the overshoot on the frequency response is already near 17%. Figure 3.5 shows a closed-loop frequency response comparison for two loop-gain conditions: maximum BWER and maximum flatness response, confirming the previous discussion. As can be seen, for the same amplifier dynamics, the loop-gain to achieve maximum flat frequency response is less than the gain to achieve maximum BWER. Similarly the closed-loop gain is larger for the maximum flat condition.

Figure 3.6 illustrates the multiple frequency peaking behavior and the associated multiple cut-off frequencies. For the considered example, the loop-gain was nearly 10 times larger than G_{max} . An important observation is the fact that the multiple peaks are indeed confined to the two bounds of (3.3). Increasing further the loop-gain or the delay produces even more peaks; the number of cut-off frequencies may become infinite.

One final observation concerning Figure 3.4 and Figure 3.6 is that in both cases a strictly linear behavior is assumed for all values of the loop-gain. In fact this linear model has restricted validity. The validity of this model is limited by two factors. First, linear models can be applied to real amplifiers assuming the operation under the small signal assumptions – since real amplifiers are composed of transistors having highly non-linear behavior, linear operation can only be assumed for small signal deviations from a given operating point. The second reason arises due to the mechanism by which oscillations are maintained in real oscillating circuits. In order to produce an oscillator using a feedback amplifier it is necessary to fulfill two requirements: the closed-loop poles must be placed on the RHP close to the imaginary axis, and have a nonlinear transfer function. A circuit in this situation is said unstable, but it can be shown that this is false for the general case. Assuming a linear behavior, the output signal due to any input perturbation (for instance, noise) contains an amplitude increasing oscillation. The presence of the nonlinear transfer function implies that the ever increasing amplitude is indeed limited (since electronic circuits have a fixed power supply, this in general traduces in limited output signal ranges). The circuit's non-linearity contradicts the increasing amplitude, forcing the oscillation to be maintained within certain bounds. The same oscillating mechanism may also appear in delayed feedback amplifiers once the oscillating conditions are reached.

This non-linear behavior also prevents the multiple cut-off phenomena of Figure 3.4 and Figure 3.6. It is very unlikely to observe such phenomena in a real circuit; however, it is still possible to observe oscillations having multiple harmonics due to the presence of the delay. Choosing to describe these phenomena with this dimension, within this discussion, is solely justified as a mathematical curiosity and also to reinforce the fact that delayed amplifiers have complex dynamics.

3.1.2 Bounded Frequency Response Lemma

The bounded frequency response lemma is a generalization of the above discussion relating equation (3.3). As shown on Figure 3.2, the existence of two bounding functions for each closed-loop frequency response can provide great insight into the design of the amplifier's frequency response. Namely, it can provide an estimate for the maximum possible closed-loop bandwidth. But perhaps more meaningful is the fact that it allows to derive meaningful guidelines for maximum flat operation achievements. These guidelines follow directly from a restricted set of necessary sufficient conditions to achieve maximal flatness in amplifiers with quasi-polynomial dynamics.

Figure 3.7 represents a unitary feedback amplifier comprising a two transfer functions: the open-loop transfer gain $A(s)$, and a controller $\beta(s)$. This representation of the feedback amplifier captures most of the properties of the traditional representation of Figure 3.1. It is easily recognized that the amplifier in Figure 3.7 converts itself into the amplifier of Figure 3.1 multiplying the closed-loop transfer function by $1/\beta(s)$. It

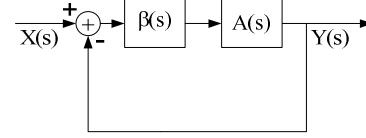


Figure 3.7 – Reference amplifier with unit feedback.

can be assumed that the controller $\beta(s)$ is a proportional controller, with fixed gain β . The loop transfer function is for this case defined by $G(s)=\beta A(s)$. Assuming that $G(s)$ fulfill the following requirements:

I – $G(s)$ can be decomposed in $G_o N(s)e^{-sL}/D(s)$, where:

$G_o=\beta A_o$ represents the loop gain, with $A_o>0$ and $\beta>0$,

L represents the total delay around the loop, with $L>0$,

$D(s)$ and $N(s)$ are polynomials with real coefficients defined as,

$$D(s) = a_0 + a_1 s + a_2 s^2 + a_3 s^3 + ..$$

$$N(s) = 1 + b_1 s + b_2 s^2 + b_3 s^3 + ..$$

where $a_0 \in \{0, 1\}$.

II – $|G(j\omega)|$ is a strictly decreasing function of ω , thus requiring that $\deg[N(s)] < \deg[D(s)]$ (necessary but not sufficient condition).

Considering now the closed-loop transfer function of the feedback amplifier of Figure 3.7 defined as,

$$H(s) = \frac{G(s)}{1 + G(s)} \quad (3.9)$$

If $G(s)$ is such that requirements I and II are true, then the squared magnitude of $H(s)$ satisfy the following lemma.

Lemma 3.1: Let $G(s)$ satisfy the assumptions I and II, then the following properties hold for the closed-loop transfer function $H(s)$:

$|H(j\omega)|^2$ is a bounded function,

$|H(j\omega)|^2$ is an asymptotically decreasing function for $\omega \geq \omega_t$.

♦ **Lemma 3.1 (proof):**

Starting with equation (3.9), $|H(j\omega)|^2$ is defined as,

$$|H(j\omega)|^2 = \frac{|G(j\omega)|^2}{|1 + G(j\omega)|^2} \quad (3.10)$$

The proof of the bounded character of $|H(j\omega)|^2$ follows directly from the modulus definition (as previously shown in (3.3)),

$$H_L(\omega) = \frac{|G(j\omega)|^2}{(1+|G(j\omega)|)^2} \leq |H(j\omega)|^2 \leq \frac{|G(j\omega)|^2}{(1-|G(j\omega)|)^2} = H_U(\omega) \quad (3.11)$$

$H_L(\omega)$ and $H_U(\omega)$ are respectively the lower and upper bounding functions of $|H(j\omega)|^2$ defined by the left and the right arms of (3.11) respectively. The second property of lemma 3.1 follows directly from the definitions of $H_L(\omega)$ and $H_U(\omega)$. Applying the definitions in I and II, the bounding functions can be rewritten as,

$$H_L(\omega) = \frac{G_o^2 |N(j\omega)|^2}{(|D(j\omega)| + G_o |N(j\omega)|)^2} \quad (3.12)$$

$$H_U(\omega) = \frac{G_o^2 |N(j\omega)|^2}{(|D(j\omega)| - G_o |N(j\omega)|)^2} \quad (3.13)$$

Since II states that $|G(j\omega)|$ is a decreasing function of ω , then it is true that $|D(j\omega)| \geq |N(j\omega)|$. Since both $|D(j\omega)|$, $|N(j\omega)|$ and G_o have by definition positive values, results that $(|D(j\omega)| + G_o |N(j\omega)|)^2 > G_o^2 |N(j\omega)|^2$, thus implying that $H_L(\omega)$ is a decreasing function of ω . The denominator of (3.13) can assume the value 0 if $|D(j\omega_b)| = G_o |N(j\omega_b)|$, thus implying that $H_U(\omega)$ is not defined in ω_b . However, for $\omega > \omega_b$, $|D(j\omega)|$ will increase faster than $G_o |N(j\omega)|$ and hence $H_U(\omega)$ is a decreasing function of ω for $\omega > \omega_b$. Since $H_U(\omega)$ is by definition continuous and decreasing for $\omega > \omega_b$, it is possible to find $\omega_k > \omega_b$ as the frequency where $H_U(\omega_k)$ meets $|H(j0)|^2$. So it is possible to state that $|H(j\omega)|^2$ is an asymptotically decreasing function of ω for $\omega \geq \omega_k$ given the fact that, $|H(j\omega)|^2$ has upper and lower bounding functions, both strictly decreasing for $\omega \geq \omega_k$. ♦

The closed-loop frequency response $H(j\omega)$ achieves maximum flat operation if $|H(j\omega)|$ is made strictly decreasing. The application of lemma 3.1 consists in the division of the frequency spectrum into two distinct sets: one above the critical frequency ω_k where $|H(j\omega)|$ is asymptotically decreasing, and another under the critical frequency ω_k where maximum flatness has to be ascertained. The following discussion will show that for feedback amplifiers with quasi-polynomial dynamics, it is very difficult to establish a strictly decreasing $|H(j\omega)|$ for all frequency spectrum. Fortunately lemma 3.1 restricts the frequency range of interest to $[0, \omega_k]$, allowing simplified maximum flatness checking procedures.

3.1.3 Maximum Bandwidth Enhancement

The BWER of a delayed feedback amplifier is defined as the ratio of the first cut-off frequency of the amplifier with delay over the cut-off frequency of the amplifier without the delay. According to the above discussion, the evaluation of the cut-off frequency for the delayed case involves the solution of a non-linear equation. The restriction of the first cut-off frequency is necessary since the complete solution can have multiple values (see appendix B 1). In order to have a clear picture of the maximum possible BWER it is necessary to consider the complete dynamics of the system and obtain plots of BWER for various delay values (assuming that both gain and open-loop poles remain fixed), which represents a major mathematical effort.

A simple alternative is to consider lemma 3.1 and the bounded nature of the closed-loop frequency response. Since the closed-loop frequency response has an upper bound, $H_U(\omega)$, it is possible to find a bound on the maximum BWER, with the ratio between cut-off frequencies of $H_U(\omega)$ and the reference amplifier without the delay. Defining the maximum cut-off frequency, ω_m , as the frequency for which,

$$H_U(\omega_m) = \frac{|H(0)|^2}{2} \quad (3.14)$$

The cut-off frequency of the reference amplifier, ω_r , is defined as the traditional 3dB attenuation point, $|H(j\omega_r)|^2 = |H(0)|^2/2$. The upper bound for the BWER follows directly from the ratio between these two cut-off frequencies,

$$BWER \leq \frac{\omega_m}{\omega_r} \quad (3.15)$$

Equation (3.14) furnishes a simplified method to evaluate the maximum BWER under the assumptions of lemma 3.1. The solutions of the maximum cut-off frequency according to (3.14) are now, the solutions of n^{th} order polynomial equation. For orders higher than the 2nd, polynomial equations have solutions which are extremely complex to put in a closed form. However, it is possible to have a qualitative picture of the behavior of the BWER functions, restricting the study for 1st and 2nd order cases.

Assuming 1st order dynamics, then $D(s) = 1 + a_1 s$. For this case the maximum BWER bound results, after simple and direct calculations in,

$$BWER_1 \leq \sqrt{\frac{1 + (3 + 2\sqrt{2})G}{1 + G}} \quad (3.16)$$

Equation (3.16) suggests that for delayed feedback amplifiers having only one pole in its internal dynamics, the maximum bandwidth improvement that can be achieved depends solely on the loop-

gain. Also, assuming that G may take very large values, according to (3.16), the maximum BWER tends to $1+2^{1/2}$. These observations are confirmed by the results depicted on Figure 3.4.

Considering the case of feedback amplifier with two poles, $D(s)$ can take the form $D(s)=1+2\xi s/\omega_n+(s/\omega_n)^2$, where ξ and ω_n are respectively the damping coefficient and the natural frequency associated to the poles of the amplifier. This representation allows a clear comprehension of the effects caused by complex poles. Following simple and direct calculations results in,

$$BWER_2 \leq \sqrt{\frac{1-2\xi^2 + \sqrt{2+G(4+2\sqrt{2}+G(3+2\sqrt{2}))} + 4\xi^2(\xi^2-1)}{1+G-2\xi^2 + \sqrt{2(1+G)^2 + 4\xi^2(\xi^2-1-G)}}} \quad (3.17)$$

Equation (3.17) shows that for the 2nd order case, the maximum BWER depends on both the loop-gain and on the damping factor. The BWER limit for large values of G is different from the previous $1+2^{1/2}$ value; in fact it is easy to see that for this case this limit is 1. The presence of a second pole destroys the bandwidth gain effect. Figure 3.8 shows the behavior of (3.17) for several values of ξ as G is swept. For $\xi > 1$, the two poles are distinct and placed on the negative real axis. When $\xi=1$, the two poles are coincident. This observation suggests that ξ can be interpreted as a measure of the poles separation. As can be seen, as ξ tends towards infinity, the 2nd order dynamics degenerate into the 1st order case, thus allowing the BWER to approach the value $1+2^{1/2}$ in the sense of (3.16). On the other hand, when the two poles come closer to each other the effect of bandwidth gain is effectively reduced. For $0 < \xi < 1$, the two poles become complex and the BWER is even smaller.

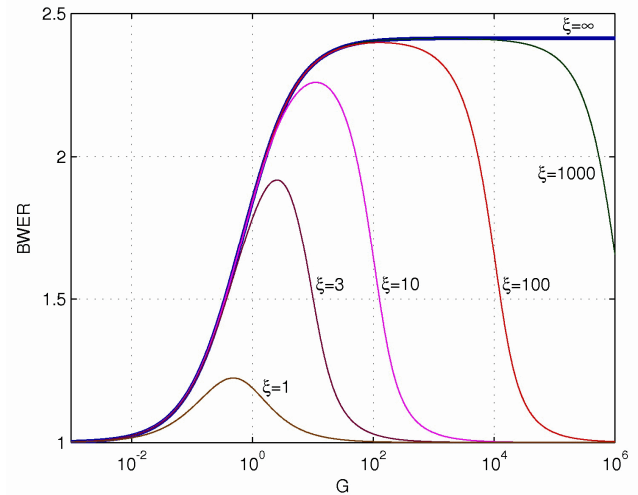


Figure 3.8 - 2nd order dynamics BWER curves.

It is not surprising that the BWER decreases as the two poles come closer together (or even become complex). For this situation the reference amplifier also experiments an increase in cut-off frequency, reaching the maximum for $\xi=2^{-1/2}$. But perhaps more important is the fact that 1st order amplifiers can provide the maximum BWER. This conclusion seems contradictory when confronted with the multi-pole feedback scheme presented on chapter 2. In chapter 2 it was shown that increasing the number of poles in a feedback amplifier improves the closed-loop bandwidth, compared to a one pole reference set-up. This effect can not be compared to the delayed feedback case because the reference set-up is different: in the delayed feedback case, the dynamics of the

open-loop are preserved except for the delay. Furthermore, in the multi-pole case, the bandwidth is greater when the poles are optimally placed having perfectly defined damping factors. For this same situation, having complex poles, the delayed feedback case presents poor performance as discussed above.

Higher order dynamics delayed feedback amplifiers have similar behaviors. The presence of more poles contributes to decrease even more the maximum achievable BWER. In order to be effective the poles must be well separated, thus implying a dominant pole behavior. Fixing the loop-gain, the maximum bandwidth gain decreases as the poles come close together. This discussion established that the presence of more than one pole in a delayed feedback contributes to decrease the maximum bandwidth for large values of the loop-gain, which in turn leads to the conclusion that 1st order delayed feedback amplifiers are the from this point of view the best choice to achieve the maximum BWER possible.

3.2 Design Constraints

This section explores the limits of operation of delayed feedback amplifiers. Theorem 3.1 (section 3.2.1) provides a method to compute necessary and sufficient conditions for maximal flat operation in delayed feedback amplifiers. Stability constraints are also analyzed (section 3.2.2), using a modification of the Hermite-Biehler theorem suitable to quasi-polynomial systems. Delayed feedback amplifiers having 1st and 2nd order dynamics are then analyzed.

3.2.1 Maximally Flat Design

It is simple to establish maximum flatness for systems without delay; it suffices to nullify the $2n-1$ first derivatives of $|H(j\omega)|^2$ for $\omega=0$ (where n is the degree of $H(s)$, see section 2.6.1 of chapter 2). However, when delays are considered, the system becomes governed by quasi-polynomial characteristic equations. In general it is impossible to nullify the $2n-1$ first derivatives of $|H(j\omega)|^2$ for these cases, since $|H(j\omega)|^2$ has non-periodic terms in $\sin(\omega L)$ and $\cos(\omega L)$ that are not all identically zero for some derivative order.

Maximum flatness can be achieved if $|H(j\omega)|^2$ can be made a decreasing function of ω at least inside some prescribed frequency set. As demonstrated in lemma 3.1, $|H(j\omega)|^2$ is at least asymptotically decreasing with ω for $\omega \geq \omega_c$. Thus it is possible to achieve maximum flatness if, condition (3.18) is taken as equality inside $[0, \omega_c[$,

$$|H(j\omega)|^2 \leq |H(0)|^2 \quad \forall \omega \in [0, \omega_c[\quad (3.18)$$

using the definitions assumed for lemma 3.1 (section 3.1.2), condition (3.18) converts into,

$$\left| D(j\omega) + G_o N(j\omega) e^{-j\omega L} \right|^2 \geq |N(j\omega)|^2 (G_o + a_0)^2 \quad \forall \omega \in [0, \omega_i[\quad (3.19)$$

since both $D(j\omega)$ and $N(j\omega)$ can be decomposed into real and imaginary polynomials,

$$D(j\omega) = d_r(\omega) + jd_i(\omega) \quad (3.20)$$

$$N(j\omega) = n_r(\omega) + jn_i(\omega) \quad (3.21)$$

condition (3.19) becomes after the variable substitution $\alpha = \omega L$,

$$\begin{aligned} d_r^2(\alpha) + d_i^2(\alpha) - a_0(a_0 + 2G_o)[n_r^2(\alpha) + n_i^2(\alpha)] &\geq \\ &\geq 2G_o[d_i(\alpha)n_r(\alpha) - d_r(\alpha)n_i(\alpha)]\sin\alpha - \\ &\quad - 2G_o[d_r(\alpha)n_r(\alpha) + d_i(\alpha)n_i(\alpha)]\cos\alpha \quad \forall \alpha \in [0, \alpha_i[\end{aligned} \quad (3.22)$$

where $\alpha_i = \omega_i L$. Defining $L(\alpha)$ and $R(\alpha)$ as,

$$L(\alpha) = d_r^2(\alpha) + d_i^2(\alpha) - a_0(a_0 + 2G_o)[n_r^2(\alpha) + n_i^2(\alpha)] \quad (3.23)$$

$$R(\alpha) = p(\alpha)\sin\alpha - q(\alpha)\cos\alpha \quad (3.24)$$

It is obvious that both $L(\alpha)$ and $R(\alpha)$ are continuous and differentiable functions, since both functions result from additions and multiplications of continuous functions. The next theorem generalizes the necessary and sufficient conditions to verify (3.22).

Theorem 3.1: *The necessary and sufficient conditions to meet condition (3.22) for a feedback system as in figure 1, meeting the requirements I and II are:*

$$\begin{aligned} G_o &\leq \frac{a_1^2 - 2a_0a_2 - a_0^2(b_1^2 - 2b_2)}{2a_2 + 2a_1(L - b_1) + a_0L(L - 2b_1) + 2a_0(b_1^2 - b_2)}, \\ L^{(3)}(\alpha) &\geq 2G_o R^{(3)}(\alpha) \quad \forall \alpha \in [0, \alpha_i[. \end{aligned}$$

◆ **Theorem 1 (proof):**

The proof of this result follows from the successive application of theorem 2.1 of chapter 2 (section 2.6.2). Applying theorem 2.1 to condition (3.22) resorts to the evaluation of the derivatives of $L(\alpha)$ and $R(\alpha)$ defined in (3.23) and (3.24) respectively. Since both $L(\alpha)$ and $R(\alpha)$ are continuous and differentiable, condition (3.22) is true if,

$$1) \left\{ \begin{array}{l} L(0) \geq 2G_o R(0) \\ L^{(1)}(\alpha) \geq 2G_o R^{(1)}(\alpha) \quad \forall \alpha \in [0, \alpha_i[\end{array} \right.$$

The first condition in 1) is easily verified, it suffices to recall the definitions in (3.23) and (3.24). Expanding the real and imaginary parts of $D(j\alpha/L)$ results in,

$$\begin{aligned}
 d_r(\alpha) &= \sum_{k=0} (-1)^k \frac{a_{2k}}{L^{2k}} \alpha^{2k} \\
 d_i(\alpha) &= \sum_{k=0} (-1)^k \frac{a_{2k+1}}{L^{2k+1}} \alpha^{2k+1}
 \end{aligned} \tag{3.25}$$

evaluating the n^{th} derivative of $d_r(\alpha)$ and $d_i(\alpha)$ at the origin,

$$\begin{aligned}
 d_r^{(2k)}(0) &= (-1)^k \frac{a_{2k}}{L^{2k}} (2k)! \\
 d_i^{(2k+1)}(0) &= (-1)^k \frac{a_{2k+1}}{L^{2k+1}} (2k+1)!
 \end{aligned} \tag{3.26}$$

where superscript (n) denotes the n^{th} α derivative. Equation (3.26) shows that both, $d_r(\alpha)$ and $d_i(\alpha)$ have zero valued n^{th} order derivatives at the origin for n odd and even respectively. The same also applies for $n_r(\alpha)$ and $n_i(\alpha)$.

$$\begin{aligned}
 n_r^{(2k)}(0) &= (-1)^k \frac{b_{2k}}{L^{2k}} (2k)! \\
 n_i^{(2k+1)}(0) &= (-1)^k \frac{b_{2k+1}}{L^{2k+1}} (2k+1)!
 \end{aligned} \tag{3.27}$$

Using (3.26) and (3.27) results,

$$L(0) = 2G_o R(0) = -2G_o a_0 \tag{3.28}$$

which verifies 1), as required. The second condition in 1) involves the first derivatives of $L(\alpha)$ and $R(\alpha)$, given by,

$$\begin{aligned}
 L^{(1)}(\alpha) &= 2 \left[d_r(\alpha) d_r^{(1)}(\alpha) + d_i(\alpha) d_i^{(1)}(\alpha) \right] - \\
 &\quad - 2a_0 [2G_o + a_0] \left[n_r(\alpha) n_r^{(1)}(\alpha) + n_i(\alpha) n_i^{(1)}(\alpha) \right]
 \end{aligned} \tag{3.29}$$

$$R^{(1)}(\alpha) = \left[p^{(1)}(\alpha) + q(\alpha) \right] \sin \alpha + \left[p(\alpha) - q^{(1)}(\alpha) \right] \cos \alpha \tag{3.30}$$

where,

$$\begin{aligned}
 p^{(1)}(\alpha) &= d_i^{(1)}(\alpha) n_r(\alpha) + d_i(\alpha) n_r^{(1)}(\alpha) - d_r^{(1)}(\alpha) n_i(\alpha) - d_r(\alpha) n_i^{(1)}(\alpha) \\
 q^{(1)}(\alpha) &= d_r^{(1)}(\alpha) n_r(\alpha) + d_r(\alpha) n_r^{(1)}(\alpha) + d_i^{(1)}(\alpha) n_i(\alpha) + d_i(\alpha) n_i^{(1)}(\alpha)
 \end{aligned} \tag{3.31}$$

Since both $L^{(1)}(\alpha)$ and $R^{(1)}(\alpha)$ are continuous and differentiable functions, it is possible to apply again theorem 2.1. The second condition of 1) is true if,

$$2) \left\{ \begin{array}{l} L^{(1)}(0) \geq 2G_o R^{(1)}(0) \\ L^{(2)}(\alpha) \geq 2G_o R^{(2)}(\alpha) \quad \forall \alpha \in [0, \alpha_i[\end{array} \right. ,$$

The first condition in 2) is easily verified since according to (3.26), (3.27), (3.29), (3.30) and (3.31), $L^{(1)}(0)=0$, and $R^{(1)}(0)=0$. The second derivatives $L^{(2)}(\alpha)$ and $R^{(2)}(\alpha)$ are given by,

$$\begin{aligned}
 L^{(2)}(\alpha) = & 2 \left[\left(d_r^{(1)}(\alpha) \right)^2 + d_r(\alpha) d_r^{(2)}(\alpha) + \left(d_i^{(1)}(\alpha) \right)^2 + \right. \\
 & \left. + d_i(\alpha) d_i^{(2)}(\alpha) \right] - a_0 (2G_o + a_0) \left[\left(n_r^{(1)}(\alpha) \right)^2 + \right. \\
 & \left. + n_r(\alpha) n_r^{(2)}(\alpha) + \left(n_i^{(1)}(\alpha) \right)^2 + n_i(\alpha) n_i^{(2)}(\alpha) \right]
 \end{aligned} \quad (3.32)$$

$$R^{(2)}(\alpha) = \left[p^{(2)}(\alpha) + 2q^{(1)}(\alpha) - p(\alpha) \right] \sin \alpha - \left[q^{(2)}(\alpha) - 2p^{(1)}(\alpha) - q(\alpha) \right] \cos \alpha \quad (3.33)$$

where,

$$\begin{aligned}
 p^{(2)}(\alpha) = & d_i^{(2)}(\alpha) n_r(\alpha) + 2d_i^{(1)}(\alpha) n_r^{(1)}(\alpha) + \\
 & + d_i(\alpha) n_r^{(2)}(\alpha) - d_r^{(2)}(\alpha) n_i(\alpha) - \\
 & - 2d_r^{(1)}(\alpha) n_i^{(1)}(\alpha) - d_r(\alpha) n_i^{(2)}(\alpha) \\
 q^{(2)}(\alpha) = & d_r^{(2)}(\alpha) n_r(\alpha) + 2d_r^{(1)}(\alpha) n_r^{(1)}(\alpha) + \\
 & + d_r(\alpha) n_r^{(2)}(\alpha) + d_i^{(2)}(\alpha) n_i(\alpha) + \\
 & + 2d_i^{(1)}(\alpha) n_i^{(1)}(\alpha) + d_i(\alpha) n_i^{(2)}(\alpha)
 \end{aligned} \quad (3.34)$$

Once again, it is possible to apply theorem 2.1 one more time, since $L^{(2)}(\alpha)$ and $R^{(2)}(\alpha)$ are both continuous and differentiable. The second condition in 2) is true if,

$$3) \left\{ \begin{array}{l} L^{(2)}(0) \geq 2G_o R^{(2)}(0) \\ L^{(3)}(\alpha) \geq 2G_o R^{(3)}(\alpha) \quad \forall \alpha \in [0, \alpha_i[\end{array} \right. ,$$

using (3.26), (3.27), (3.32), (3.33) and (3.34) results in,

$$\begin{aligned}
 L^{(2)}(0) = & \frac{2}{L^2} \left[a_1^2 - 2a_0 a_2 - a_0^2 (b_1^2 - 2b_2) \right] - 4 \frac{G_o a_0}{L^2} (b_1^2 - 2b_2) \\
 R^{(2)}(0) = & -q^{(2)}(0) + 2p^{(1)}(0) + q(0) = \\
 = & a_0 \left(1 - \frac{2b_1}{L} + \frac{2b_2}{L^2} \right) + \frac{2a_1}{L} \left(1 - \frac{b_1}{L} \right) + \frac{2a_2}{L^2}
 \end{aligned} \quad (3.35)$$

Combining $L^{(2)}(0)$ and $R^{(2)}(0)$ gives the necessary condition of theorem 1.

$$G_o \leq \frac{a_1^2 - 2a_0 a_2 - a_0^2 (b_1^2 - 2b_2)}{2a_2 + 2a_1 (L - b_1) + a_0 L (L - 2b_1) + 2a_0 (b_1^2 - b_2)} \quad (3.36)$$

The sufficient condition of theorem 1 follows from the third derivative test in 3), thus completing the proof. \blacklozenge

At a first glance it seems possible to proceed with these derivative tests, however, this is not so. A careful observation reveals that $L^{(u)}(\alpha)$ becomes identically zero for a well defined derivative order, since $L(\alpha)$ is a polynomial of known order. The same is not verified for $R^{(u)}(\alpha)$. Inspecting

the first two derivatives of $R(\alpha)$ reveals that the original order of $p(\alpha)$ and $q(\alpha)$ is preserved throughout the whole procedure. So it is very likely to happen that $L^{(u)}(\alpha) \geq 2k_p R^{(u)}(\alpha)$ will fail to verify for $u > 3$, because $L^{(u)}(\alpha)$ will progressively increase slower while $R^{(u)}(\alpha)$ will progressively oscillate with larger amplitudes.

3.2.2 Stability Limits

There are several possible methods to study the stability restrictions of feedback time-delay systems. This analysis can be divided into three main types: time domain, frequency domain and state-space methods. Due to the nature of the present problem, bandwidth optimization in feedback amplifiers, the most immediate methods to study stability use a frequency domain representation. Among these frequency domain methods, several techniques can be applied, namely: the Nyquist stability criterion [257, 283]; analytical root-locus techniques [133-135], and generalizations of the classical Hermite-Biehler theorem to quasi-polynomials [132, 149-152, 154, 155, 285] among others [282]. Due to its concise formulation, the modified Hermite-Biehler theorem allows a complete description of the stability restrictions for these kinds of systems. In order to provide a short and comprehensive description of this theorem, some basic notions must be presented. A complete and exhaustive exploration of this material can be found in [282, 283, 285].

3.2.2.1 Preliminary Definitions

A general characteristic quasi-polynomial is defined as a function containing both polynomial terms and complex exponential terms. Mathematically this is expressed as,

$$\delta(s) = d(s) + n_1(s)e^{-sL_1} + n_2(s)e^{-sL_2} + \dots + n_m(s)e^{-sL_m} \quad (3.37)$$

where $d(s)$ and $n_i(s)$ are real coefficient polynomials, L_i are all real non-negative.

Several types of systems can be described by this general representation. If all the L_i terms can be related by $L_i = \rho_i L_1$ with all ρ_i being non-negative integers, then the system is said to exhibit commensurate delays. For this case the quasi-polynomial in (3.37) takes the form of a two variable polynomial $f(s, e^{-s})$, for which there are available methods to perform stability analysis. When ρ_i are not all integers, the system is said to exhibit a incommensurate delay nature. For this kind of systems, the problem of establishing stability may turn to be very complex [282].

Considering only commensurate systems, assuming that $0 < L_1 < L_2 < \dots < L_m$, it is possible to consider $\delta^*(s) = \delta(s)e^{sL_m}$ instead of (1.37). Making these transformations, equation (3.37) becomes,

$$\delta^*(s) = d(s)e^{sL_m} + n_1(s)e^{s(L_m-L_1)} + n_2(s)e^{s(L_m-L_2)} + \dots + n_m(s) \quad (3.38)$$

Note that since $e^{sL_m} > 0$ the zeros of $\delta^*(s)$ and $\delta(s)$ are exactly the same. Now substituting $t = e^{sL_1}$ in

(3.38) $\delta^*(s)$ takes the form of a double summation on the variables s and t ,

$$\delta^*(s) = f(s, t) = \sum_{h=0}^M \sum_{k=0}^N a_{hk} s^h t^k \quad (3.39)$$

If a_{MN} exists and has nonzero value, then $f(s, t)$ is said to have a principal term. The presence or not of a principal term further divides the system classification. Thus, systems having commensurate delay nature can be classified into three further subtypes: 1) retarded-type systems (or delay-type), when $\deg[d(s)] > \deg[n_i(s)]$ for all i - this means that the principal term is attached to the highest powers of both s and t alone; 2) neutral-type systems, when $\deg[d(s)] = \deg[n_i(s)]$ for some i - this means that there are at least two terms having the same power of s , one being the principal term; and 3) Forestall-type systems, when $\deg[d(s)] < \deg[n_i(s)]$ for some i - this means that there is no principal term. Forestall-type systems are always unstable. Pontryagin established that for this case there exist always an unbound number of roots in the RHP [132, 285].

In general, stability of retarded and neutral type quasi-polynomials is defined in a similar manner. For a retarded type quasi-polynomial, stability is ascertained if all its roots have negative real parts, that is, if s is a root of the quasi-polynomial, then $\text{Re}[s] < 0$. Neutral type quasi-polynomials, on the other hand require a stringent argument on its roots. For a neutral type quasi-polynomial, stability is ascertained if all its roots have real parts satisfying the condition, $\text{Re}[s] < -\sigma$, for some $\sigma > 0$ and real. This restriction is necessary because neutral systems may exhibit root chains on the LHP that may converge asymptotically to the imaginary axis.

3.2.2.2 Available Results on the Stability of Quasi-Polynomials

Pontryagin published several results on the zeros of commensurate quasi-polynomials of all the referred types [132]. One of these results plays an important role on the stability assessment of retarded type quasi-polynomials. These results form the foundations that allowed Silva *et al* to extend the classical Hermite-Biehler theorem to quasi-polynomial cases [149, 285]. Consider $\delta^*(s) = f(s, e^s)$ decomposed into real and imaginary parts,

$$\delta^*(s) = \delta_r(s) + j\delta_i(s) \quad (3.40)$$

It is clear that both $\delta_r(s)$ and $\delta_i(s)$ are functions containing polynomial, sine and cosine terms. Consider now the restriction of s to the imaginary axis, $j\omega$,

Theorem 3.2: Let M and N denote the highest powers of s and e^s , respectively in $\delta^*(s)$. Let η be an appropriate constant such that the coefficients of the terms of highest degree in $\delta_r(\omega)$ and $\delta_i(\omega)$ do not vanish at $\omega = \eta$. Then for the equations $\delta_r(\omega) = 0$ and $\delta_i(\omega) = 0$ to have only real roots, it is

necessary and sufficient that in each of the intervals

$$-2l\pi + \eta \leq \omega \leq 2l\pi + \eta, \quad l = l_o, l_o + 1, l_o + 2, \dots$$

$\delta_r(\omega)$ and $\delta_i(\omega)$ have exactly $4lN+M$ real roots for a sufficiently large l_o .

The classical Hermite-Biehler theorem establishes the necessary and sufficient conditions for a real coefficient polynomial to be Hurwitz stable [283, 285]¹³. In its original form, the Hermite-Biehler is based on one important property that the real and imaginary components of a stable polynomial must fulfill on the imaginary axis: the interlacing property (an exploration of the properties of Hurwitz stable polynomials can be found in [138, 142, 146, 283, 285]). Extensions of the classical Hermite-Biehler theorem to include polynomials with roots on the imaginary axis and also on the RHP where proposed by Ho *et al* in [144, 147]).

Using theorem 3.3, Silva *et al* demonstrated that, stable quasi-polynomials verify the interlacing property [149]. In this sense, it is important to define the root interlacing property. Consider the real and imaginary partition of $\delta^*(s)$ stated on (3.40), for which, $\omega_{r1}, \omega_{r2}, \omega_{r3}, \dots$, represent the roots of $\delta_r(\omega)$, and $\omega_{i1}, \omega_{i2}, \omega_{i3}, \dots$, represent the roots of $\delta_i(\omega)$. The roots of $\delta_r(\omega)$ and $\delta_i(\omega)$ obey the interlacing property if,

$$\omega_{r1} < \omega_{i1} < \omega_{r2} < \omega_{i2} < \dots$$

Combining theorem 3.2 with the interlacing property definition, the modified Hermite-Biehler theorem for quasi-polynomials becomes,

Theorem 3.3: Let $\delta^*(s) = f(s, e^s)$ be a quasi-polynomial representing a system with commensurate delays, furthermore, let $f(s, t)$ be a polynomial in s and t with principal term, then consider the form,

$$\delta^*(j\omega) = \delta_r(\omega) + j\delta_i(\omega)$$

where $\delta_r(\omega)$ and $\delta_i(\omega)$ are respectively the real and imaginary parts of $\delta^*(j\omega)$. $\delta^*(s)$ is stable if and only if

$\delta_r(\omega)$ and $\delta_i(\omega)$ have only simple, real roots and these interlace.

$\delta_i'(\omega_o)\delta_r(\omega_o) - \delta_r'(\omega_o)\delta_i(\omega_o) > 0$, for some $\omega_o \in]-\infty, \infty[$.

Theorems 3.2 and 3.3 form an adequate framework to perform stability analysis in commensurate quasi-polynomials systems. According to the lemma 3.1 and its assumptions, delayed feedback amplifiers are a particular set of commensurate quasi-polynomial systems, for which these results do apply.

¹³ Hurwitz stable polynomials are polynomials having all its roots in the open LHP.

Silva *et al* applied successfully these results to first and second order feedback systems with one delay [150, 151]. Delayed feedback amplifiers having first and second order dynamics fall into this category. Unfortunately, Silva *et al* results are stated in a pure mathematical form which is difficult to apply directly to delayed feedback analysis. Theorems 3.4 and 3.5 present suitable translations for this case.

Theorem 3.4: A delayed feedback amplifier with loop transfer function $G(s)$ given by,

$$G(s) = G_o \frac{e^{-sL}}{1 + a_1 s}$$

where G_o , a_1 and L are all positive, is Hurwitz stable for G_o values satisfying the following restriction:

$$0 < G_o < \sqrt{1 + \frac{a_1^2}{L^2} z_1^2}$$

where z_1 is the solution of $\tan(z) = -\frac{a_1}{L} z$ in the interval $] \pi/2, \pi[$.

Theorem 3.5: A delayed feedback amplifier with loop transfer function $G(s)$ given by

$$G(s) = G_o \frac{e^{-sL}}{1 + 2\xi s/\omega_n + (s/\omega_n)^2}$$

where G_o , L positive and Hurwitz stable denominator, is Hurwitz stable for G_o values satisfying the following restriction:

$$0 < G_o < \frac{\sqrt{(\omega_n L)^4 + 2G_2 (\omega_n L)^2 z_1^2 + z_1^4}}{(\omega_n L)^2}$$

where z_1 is the solution of $\cot(z) = \left[z^2 - (\omega_n L)^2 \right] / (2\xi \omega_n L z)$ in the interval $]0, \pi[$.

Appendix B 2 and B 3 provide a short proof of these two theorems.

As stated on theorems 3.4 and 3.5, the stability bounds for 1st and 2nd order delayed feedback systems depend on the solution of a non-linear equation. This is in general true for systems having arbitrary orders. However, for some special kinds of delayed feedback amplifiers it is also possible to obtain in closed forms the roots of the quasi-polynomial characteristic equation. These solutions are expressed as branches of the Lambert W function [148, 153, 156]. However these are not general solutions for general delayed feedback amplifiers, and cannot provide adequate stability bounds.

3.3 Delayed Feedback Amplifiers Types

Section 3.1 described the constitution of the delayed feedback amplifier. According to the discussion, a delayed feedback amplifiers consists of a feedback amplifier having a delay element inside its feedback loop. This description is very ambiguous, since the location of the delay element can be anywhere inside the feedback loop. Furthermore, the same discussion also presented two kinds of delay elements. According to their mathematical description, they can be: of pure delay type – when represented by a complex exponential; or all-pass type – when represented by a suitable approximation of the complex exponential. Also relevant to the design is the number of poles and zeros contained in the loop transfer function, since this impacts the maximum achievable BWER.

A suitable classification of delayed feedback amplifiers comprises three basic classes: pure delay, all-pass delay and mixed delay. Figure 3.9 represents these three classes. The symbol $\Delta(s)$ represents a all-pass delay transfer function. All of three classes comprise a forward amplifier $A(s)$ and a constant feedback factor, β . Their difference is the presence of different delay elements on the feedback loop. For now, it will be assumed that the total delay around the loop is concentrated on the feedback path. However, it should be clear that this is not a suitable representation of real delayed feedback amplifiers, since all the elements that compose the amplifier are bounded by physical laws and thus, all of them have its own time delay contribution. The effects of this “distributed” nature of the delay will be explored on section 3.5.

The application of stability and maximal flat constraints to the design of delayed feedback amplifiers is dependent on the type of dynamic model adopted for the forward amplifier. In this work, the common approaches of dominant pole and of two-pole amplifiers where adopted to illustrate the concept. Application of these results to other kinds of delayed feedback amplifiers is accomplished using similar procedures, with increased mathematical complexity.

3.3.1 1st Order Delayed Feedback Amplifiers

The following theorem verifies the maximum flatness constraints for a delayed feedback amplifier having only one pole in its forward transfer function. This representation captures the nature of dominant pole amplifiers. According to assumptions I and II of section 3.1.2, the internal dynamics are represented by, $N(s)=1$, and $D(s)=1+a_1s$. With these assumptions $G(s)$ is given by,

$$G(s) = G_o \frac{e^{-sL}}{1+a_1s} \quad (3.41)$$

where $G_o=\beta A_o$ is the loop-gain. From the maximum flatness perspective, a_i can be either positive or

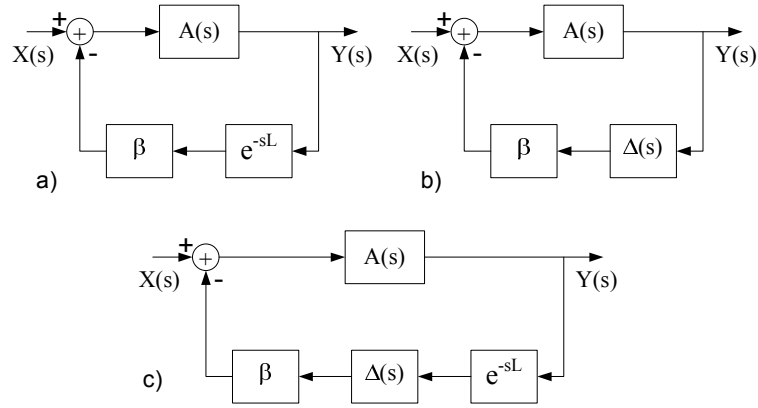


Figure 3.9 - Delayed feedback amplifier classes: a) pure delay; b) all-pass delay; and c) mixed delay.

negative, but the resulting $|G(j\omega)|$ is for both cases strictly decreasing with ω . However, for stability reasons it will be assumed that $a_1 > 0$.

Theorem 3.6: A delayed feedback amplifier with loop transfer function $G(s)$ given by (3.41) where G_o , a_1 and L are all positive achieves maximum flatness if

$$G_o = \frac{a_1^2}{L(L + 2a_1)}$$

Under this condition the amplifier is also Hurwitz stable.

◆ **Theorem 3.6 (proof):**

Applying the results of theorem 3.1 (section 3.2.1) the necessary condition in order to have $|H(j\omega)|^2$ monotonically decreasing inside $[0, \omega_t]$ is given by.

$$G_o \leq \frac{a_1^2}{L(L + 2a_1)} \quad (3.42)$$

Taking this result as an equality $|H(j\omega)|^2$ will be monotonically decreasing in a maximum sense inside $[0, \omega_t]$; or equivalently, $|H(j\omega)|^2$ is maximally flat, since for $\omega \geq \omega_t$ $|H(j\omega)|^2$ is asymptotically decreasing. The sufficiency of this result follows from the second condition in theorem 3.1. Forming $L^{(3)}(\alpha)$ and $R^{(3)}(\alpha)$ results in,

$$L^{(3)}(\alpha) = 0 \geq -\left(1 + 3\frac{a_1}{L}\right)\sin\alpha - \frac{a_1}{L}\alpha\cos\alpha = R^{(3)}(\alpha) \quad \forall \alpha \in [0, \alpha_t] \quad (3.43)$$

From (3.43), it is clear that $R^{(3)}(\alpha)$ is negative for small values of α above 0; since $R^{(3)}(0) = 0$ it must also be decreasing. This observation makes possible to verify (3.43) for a small set from 0 to α_0 , where α_0 represents the first zero of $R^{(3)}(\alpha)$ in $]0, \infty[$. The sufficiency condition is met for all possible values of a_1 and L if α_0 is larger than the critical frequency of lemma 3.1, α_t , defined from $H_U(\omega_t) = |H(j0)|^2$,

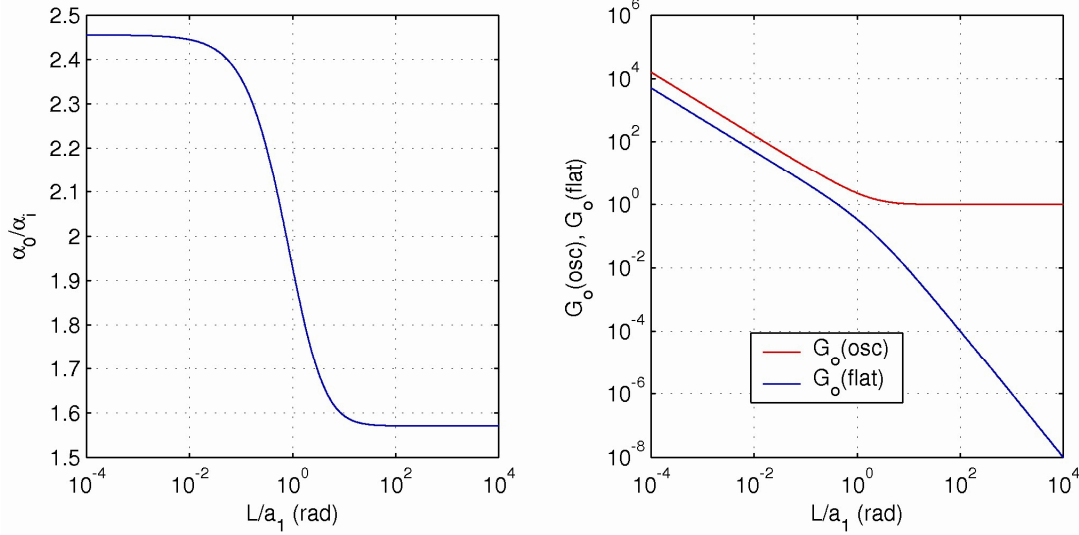


Figure 3.10 – Maximum flatness and stability constraints for a 1st order delayed feedback amplifier.

$$\alpha_i = \frac{2G_o L}{a_1} \sqrt{1 + \frac{1}{G_o}} \quad (3.44)$$

The explicit solution of $R^{(3)}(\alpha)=0$ is difficult to express in a closed form. However, it is simple to find its numerical solution for several values of the pole-delay product, L/a_1 . Figure 3.10 depicts the ratio between α_0 and α_i for practical values of $L/a_1 > 0$. $L/a_1 = \infty$ represents a trivial case with $a_1 = 0$, where maximum flatness naturally holds. Figure 3.10 shows that indeed, α_0 is larger than α_i . This completes the proof.

The stability proof follows from theorem 3.4 (section 3.2.2.2). Once again it is difficult to express in a closed form the values of z_I , as required by theorem 3.4. An alternative solution is to use numerical simulation to study the stability bound as required. Figure 3.10 shows the comparison between the values of G_o needed to achieve the flatness condition, $G_o(flat)$, and the stability limit $G_o(osc)$. This was done for the same values of the pole-delay product as above. Figure 3.10 shows the flatness condition results in loop-gain values inferior to the stability limit, so the system is stable under the maximum flatness condition. This completes the proof. ♦

3.3.2 2nd Order Delayed Feedback Amplifiers

Assuming now that the forward amplifier has a two-pole dynamic, with $N(s)=1$, $D(s)=1+2\xi s/\omega_n+(s/\omega_n)^2$, G_o and L both positive and $D(s)$ Hurwitz stable, the loop transfer function is defined by,

$$G(s) = G_o \frac{e^{-sL}}{1 + 2\xi s/\omega_n + (s/\omega_n)^2} \quad (3.45)$$

With these assumptions, it is easily seen that $G(s)$ satisfies I. Requirement II is satisfied if ω_n

and ξ are such that $|D(j\omega)|$ is increasing with ω , which is assured with $\xi > 0.707$. The following theorem verifies the necessary and sufficient conditions to achieve maximum flatness for all possible values of ω_n , ξ and L .

Theorem 3.7: A delayed feedback amplifier with $G(s)$ defined in (3.45) with G_o , L positive and $D(s)$ Hurwitz stable achieves the maximum flatness if

$$G_o = \frac{2(2\xi^2 - 1)}{(\omega_n L)^2 + 4\xi\omega_n L + 2}$$

Under this condition the amplifier is also Hurwitz stable.

♦ **Theorem 3.7 (proof):**

Applying theorem 3.1, the necessary condition is,

$$G_o \leq \frac{2(2\xi^2 - 1)}{(\omega_n L)^2 + 4\xi\omega_n L + 2} \quad (3.46)$$

As stated previously, requirement II of lemma 3.1 implies that the maximum allowed value given by (3.46) is positive. The sufficiency of (3.46) is ascertained if $L^{(3)}(\alpha) \geq 2G_o R^{(3)}(\alpha)$ holds inside $[0, \alpha_i]$. Constructing $L^{(3)}(\alpha)$ and $R^{(3)}(\alpha)$ leads to,

$$12\alpha \geq -2G_o (\omega_n L)^2 (\xi\omega_n L + 3)\alpha \cos \alpha - \left(6 + 6\xi\omega_n L + (\omega_n L)^2 - \alpha^2\right) \sin \alpha \quad \forall \alpha \in [0, \alpha_i] \quad (3.47)$$

As before, since $R^{(3)}(\alpha)$ is negative and decreasing for small values of α above 0, while $L^{(3)}(\alpha)$ is positive and increasing, it is possible to meet the sufficiency condition. Let α_0 represents the first zero of $L^{(3)}(\alpha) = 2G_o R^{(3)}(\alpha)$ inside $]0, \infty[$, when G_o is taken as the maximum value defined by (3.46). Sufficiency results if α_0 is larger than the critical frequency α_i , given by

$$\alpha_i = \omega_n L \sqrt{G_2 \sqrt{1 + 4\left(\frac{G_o}{G_2}\right)^2 \left(1 + \frac{1}{G_o}\right)} - G_2} \quad (3.48)$$

where G_2 represents the limit of G_o when L becomes zero.

$$G_2 = \lim_{L \rightarrow 0} G_o|_{flat} = 2\xi^2 - 1 \quad (3.49)$$

Finding α_0 is again feasible using numerical simulation, using $\omega_n L$ and ξ as variable parameters. Figure 3.11 shows the result of the comparison of α_0 and α_i for $\xi > 0.707$; the condition $\alpha_0 \geq \alpha_i$ is met for all values of $\omega_n L > 0$. When $\xi = 0.707$ the open-loop amplifier is maximally flat. It is not possible to attain maximum flatness under closed-loop operation for this situation, since the amplifier has already maximum flat response under open-loop conditions. This is supported by the result of (3.46)

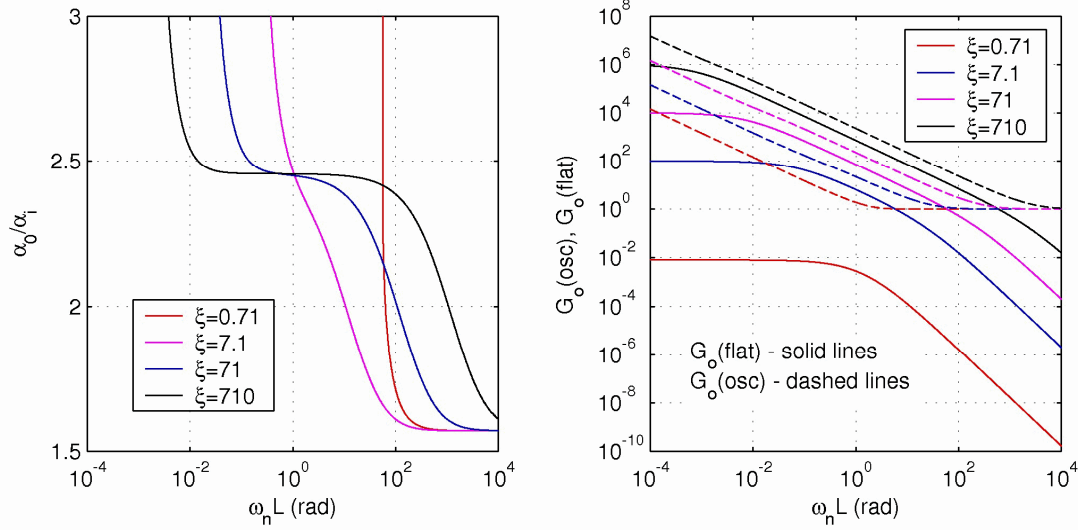


Figure 3.11 - Maximum flatness and stability constraints for a 2nd order delayed feedback amplifier.

$G_o=0$. This concludes the sufficiency proof.

Finally the stability proof follows from theorem 3.5 (3.2.2.2). Graphically Figure 3.11 shows the comparison between the values of G_o needed to achieve the flatness condition $G_o(flat)$, and the stability limit $G_o(osc)$. This was done for positive values of $\omega_n L > 0$ and several values of $\xi > 0.707$ as before. Figure 3.11 shows the flatness condition results in loop-gain values inferior to the stability limit, so the system is stable under the maximum flatness condition. This completes the proof. ♦

3.3.3 Delayed Feedback Amplifiers with All-Pass Delays

Pure delays are represented by exponential functions of the complex frequency variable. This representation is also a possible form for the transfer function of an all-pass filter, that is, a filter able to act on the phase components of a signal, without changing its amplitude. A suitable method to approximate this type of transfer functions is to use Padé approximants. As shall be discussed in chapter 4, these approximations are particularly suited to implementations using active devices. Padé's approximants are rational approximations of functions represented by their power series expansion [268]. Given a generic function $f(s)$, where s may be take complex values, its power series can be obtained using a Maclaurin expansion, such as,

$$f(s) = \sum_{i=0}^{\infty} c_i s^i \quad (3.50)$$

A Padé approximant of $f(s)$ is a rational function of the form,

$$R_p^k(s) = \frac{a_0 + a_1 s + \dots + a_k s^k}{b_0 + b_1 s + \dots + b_p s^p} \quad (3.51)$$

which has a Maclaurin expansion that agrees as far as possible with (3.50). Normally, b_0 is set to

1, leaving a total of $k+p+1$ coefficients to determine. This number suggest that $R_p^k(s)$ is ought to fit the power series of $f(s)$ through the orders 0 till $k+p$, that is,

$$f(s) = \sum_{i=0}^{\infty} c_i s^i = \frac{a_0 + a_1 s + \dots + a_k s^k}{b_0 + b_1 s + \dots + b_p s^p} + O(s^{k+p+1}) \quad (3.52)$$

where $O(s^{k+p+1})$ represents the truncated factors of the original series expansion. Coefficients a_j e b_k can be easily computed, cross multiplying the denominator of the fraction in (3.52) by the first $k+p$ terms of the Maclaurin expansion of $f(s)$ and equating the equal powers. Two sets of equations give the values of b_l and a_j as function of coefficients c_i .

3.3.3.1 All-Pass Delays

Particularly relevant for the present study are Padé approximants of the complex exponential. As will be demonstrated on the next chapter, these approximants are useful from a circuit design point of view as they allow simple and accurate delay element realizations. Padé approximants can also be used to study stability properties of time-delay systems as stressed out in [282, 285]. However, the approximate nature of such techniques leads to stability bounds that are somehow optimistic [285]. This is easily understandable, since Padé approximants of the complex exponential provide system dynamics that are only valid for restricted frequencies (an obvious observation is the fact that the approximated characteristic equation has only a finite number of roots directly related to the degree of the approximant). Nevertheless, from the implementation point of view, Padé approximations on delayed feedback amplifiers can represent a valuable tool.

In lemma 3.1 it was established that, from the delayed feedback perspective, the optimizing elements should only affect the phase response of the loop-transfer function. As it was previously discussed, all-pass filters having complex exponential representations provide one such possibility. The usage of Padé approximants can only be considered under the same assumption. Fortunately, it is established that Padé approximants of the complex exponential having same numerator/denominator degrees are all-pass transfer functions. In fact, it was shown that for these case it is possible to write [256, 268],

$$e^{-sL} \approx R_k^k(s) = \frac{P_k(-s)}{P_k(s)} \quad (3.53)$$

where $P_k(s)$ is an k^{th} degree Hurwitz stable polynomial. The roots of $P_k(-s)$ are symmetric images of the roots of $P_k(s)$. Depending on the order of the approximant, k , $P_k(s)$ takes the form,

$$P_k(s) = \sum_{l=1}^k \frac{(2k-l)!}{(k-l)!} x^l \quad (3.54)$$

Table 3.1 – Design constraints for delayed feedback amplifiers with a simulated delay of $2T$ seconds.

n	Maximum Flatness	Stability
1	$G_o = \frac{a_1^2}{4T(T+a_1)}$	$0 < G_o < 1 + \frac{a_1}{T}$
2	$G_o = \frac{2\xi^2 - 1}{1 + 2\omega_n T(\omega_n T + 2\xi)}$	$0 < G_o < 1 + \frac{\xi + \omega_n T(2\xi^2 - 1)}{\omega_n T(1 + \xi\omega_n T)}$

Increasing the order of the approximant improves the quality of the stability prediction, as stated in [285]. However, this is not efficient from a design perspective, since increasing k increases the complexity of the design. A more efficient approach is to consider the simplest form of an all-pass delay. According to (3.53) and (3.54), taking $k=1$ produces,

$$e^{-sT} \approx \Delta(s) = \frac{1-sT}{1+sT} \quad (3.55)$$

Padé approximants with the form of (3.55) have been used to implement accurate delay elements in electronic circuits [207, 298-299]. Implementation details will be considered on the next chapter, but for now it is important to explore the design constraints on this kind of amplifiers.

For delayed feedback amplifiers having all-pass delays, both maximal flat and stability constraints follow in a more direct fashion. Maximal flatness follows from equation (2.51) of chapter 2, since for this case the number of derivatives of $|H(j\omega)|^2$ is finite. Accessing stability restrictions is also easier; for this case it is possible to apply the Routh-Hurwitz criterion [257, 283]. Table 3.1 summarizes these design constraints for 1st and 2nd order delayed feedback amplifiers having an all-pass delay element with a net delay of $2T$ seconds. It is interesting to notice that the maximum flatness gain has exactly the same value predicted for the pure delay case by equations (3.42) and (3.46), when L is replaced by $2T$.

3.3.3.2 Mixed Delays

A more general class of delayed feedback amplifiers comprises both types of delay elements, pure delays and all-pass delays, as shown on Figure 3.9c. This class of amplifiers has a much more complicated dynamical description. The loop transfer function to consider has the form (assuming a first order all-pass transfer function),

$$G(s) = \frac{G_o}{1+a_1s} \frac{1-sT}{1+sT} e^{-sL} \quad (3.56)$$

The effective delay in such a case can be evaluated summing the pure delay and simulated delay contributions, or, $L_{eff}=L+2T$. The following theorem establishes the conditions to attain maximum flat operation for a 1st order delayed feedback amplifier with mixed delays.

Theorem 3.8: A delayed feedback amplifier with loop transfer function $G(s)$ given as in (3.56) where k_p , a_1 , T and L are all positive, achieves maximum flatness if

$$G_o = \frac{a_1^2}{(L+2T)(2a_1+L+2T)}$$

♦ **Theorem 3.8 (proof):**

Applying the results of theorem 3.1, the necessary condition in order to have $|H(j\omega)|^2$ monotonically decreasing inside $[0, \omega_i]$ is given by,

$$G_o \leq \frac{a_1^2}{(L+2T)(2a_1+L+2T)} \quad (3.57)$$

Taking this result as an equality $|H(j\omega)|^2$ will be monotonically decreasing in a maximum sense inside $[0, \omega_i]$; or equivalently, $|H(j\omega)|^2$ is maximally flat, since for $\omega \geq \omega_i$ $|H(j\omega)|^2$ is asymptotically decreasing. The sufficiency of this result follows from the second condition in theorem 3.1. Forming $L^{(3)}(\alpha)$ and $R^{(3)}(\alpha)$ results in,

$$\begin{aligned} 12u_1^2u_2^2\alpha \geq -G_o \left\{ \left[u_1 + 2u_2(1+6u_1) + 6u_2^2(1+3u_1) - u_1u_2^2\alpha^2 \right] \alpha \cos \alpha \right. \\ \left. + \left[1 + 3u_1 + 6u_2(1+2u_1) + 6u_2^2(1+u_1) - (2u_1u_2 + u_2^2(1+9u_1))\alpha^2 \right] \sin \alpha \right\} \quad \forall \alpha \in [0, \alpha_i] \end{aligned} \quad (3.58)$$

where $u_1 = a_1/L$ and $u_2 = T/L$. As before, the right side of (3.58) is negative and decreasing for small values of α above 0. Let α_0 represent the first zero of (3.58) in $]0, \infty[$. If α_0 is larger than α_i (given by (3.48), since $|1+j\omega T| = |1-j\omega T|$) the sufficiency condition is met for all possible values of a_1 , T and L . The explicit solution of (3.58) is difficult to express in a closed form. However, it is possible to find numerical solutions for several values of the parameter u_1 and u_2 . In order to provide a direct comparison between delayed feedback amplifiers with mixed delays and delayed feedback amplifiers having other types of delays, L and T in parameters u_1 and u_2 of equation (3.58) should be replaced by the total delay, L_{eff} . This procedure allows a direct comparison between the three cases for the same pole-delay product. Expressing L and T as functions of L_{eff} and a modifying parameter ϕ , results,

$$\begin{aligned} u_1 = \frac{a_1}{L} = \frac{1}{\phi} \frac{a_1}{L_{eff}} \\ u_2 = \frac{T}{L} = \frac{1-\phi}{2\phi} \end{aligned} \quad (3.59)$$

where ϕ represents the percentage of pure delay to the effective delay, $\phi = L/L_{eff}$. In this sense ϕ can be seen as a modifying parameter, quantifying the behavior distinction between the pure delay and the all-pass delay cases.

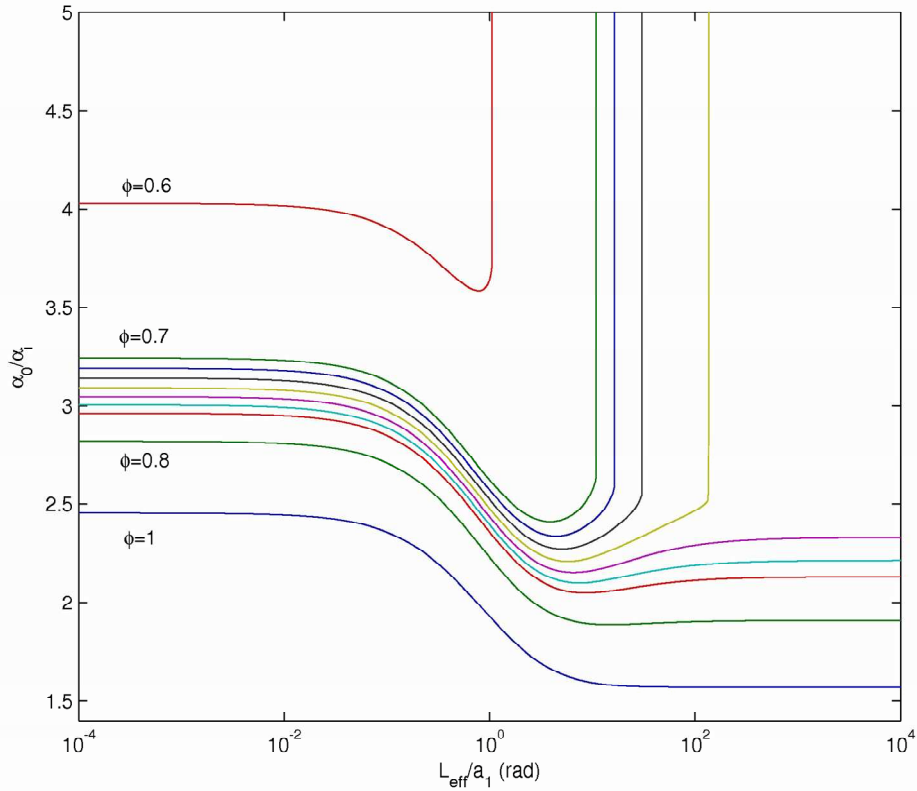
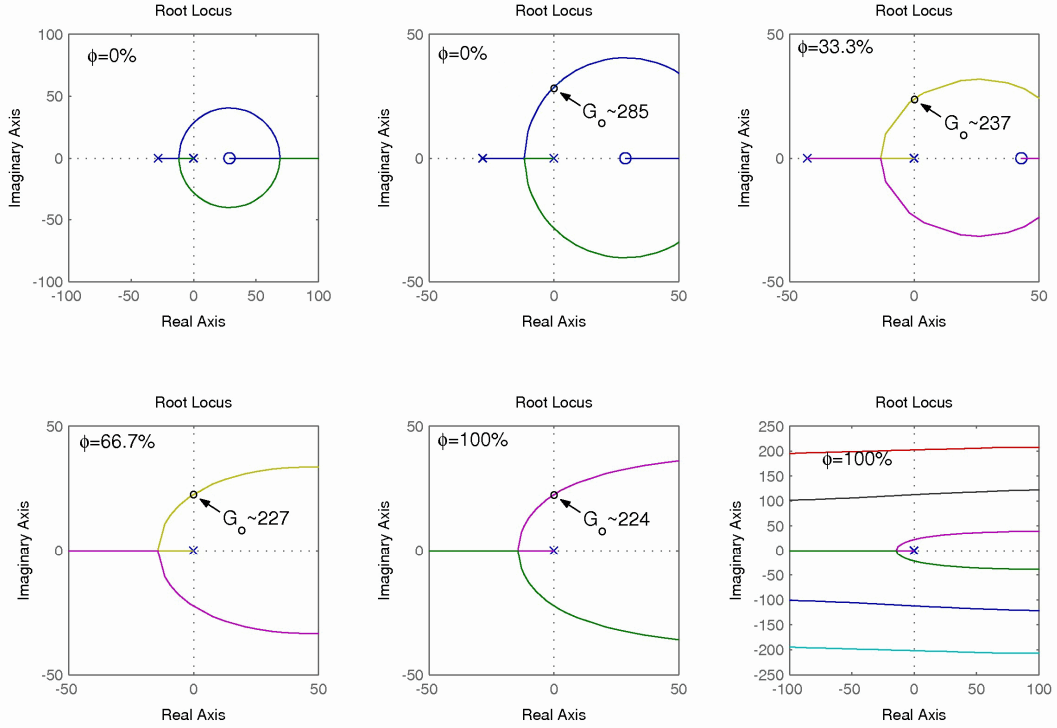


Figure 3.12 - Maximum flatness constraints for a 1st order delayed feedback amplifier with mixed delays.

Figure 3.12 depicts a comparison between α_0 and α_i for several values of the effective pole-delay product, L_{eff}/a_1 , and ϕ . Figure 3.12 shows that α_0 is always larger than α_i . Also shown is the presence of a different behavior for values of ϕ less than ~ 0.7 . The reason for this lies on the fact that the first zero of (3.58) may exhibit a discontinuous nature depending on both parameters. Since the right side of (3.58) is an amplitude increasing oscillating function, the multiple crossings against $12u_1^2 u_2^2 \alpha$ may not always lie on the first semi-period. For smaller values of ϕ , α_0 becomes even larger. This increase is in complete agreement with the result for the simulated delay case; since smaller values of ϕ represent an effective proximity with this case, for which there is no restriction on α_0 . This completes the proof. ♦

It is interesting to notice that, as in the case of all-pass delays, the maximum flat gain restriction of equation (3.57) has the same form of (3.42), taking the same effective delay $(L+2T)$. This suggests that delayed feedback amplifiers with mixed delays have an intermediate behavior between the pure delay and all-pass delay situations. In fact taking $\phi=1$ or $\phi=0$ results in the two extreme cases, pure delays and all-pass delays, respectively. From the stability point of view this represents a useful simplification, since both stability restrictions of the previous cases can be used to ascertain the stability of (3.57).

Theorem 3.9 does not provide any insight on the stability of condition (3.57). A complete


 Figure 3.13 – Root-locus for mixed delay amplifiers for several values of ϕ .

demonstration of stability implies the application of the modified Hermite-Biehler theorem as in the previous cases. However due to the presence of the all-pass delay element, the verification procedure becomes extremely difficult. As shown in appendix B 4, this procedure depends on the solution of a system of non-linear equations both dependent on the parameters u_1 and u_2 .

Another possibility to illustrate stability is to consider the intermediate nature between the pure delay and the all-pass delay cases for which the stability restrictions were previously expressed. A better illustration of the intermediate nature of this kind of amplifiers can be adequately represented using root-locus techniques. Figure 3.13 shows the root-locus modifications as ϕ varies between 0 and 1, for a mixed delay amplifier having $a_1=10\text{s/rad}$ and $L_{\text{eff}}=70\text{ms}$. The first two figures on the top left represent the extreme case with $\phi=0$, all-pass delay. For this case the stability restriction on the loop-gain has the value of ~ 285 , as given from Table 3.1. The other extreme case, $\phi=1$, pure delay is presented on the bottom right of Figure 3.13. For this case the stability restriction imposes a smaller loop-gain value of ~ 224 (also in accordance to the results of theorem 3.6). For values of ϕ between 0 and 1 the loop-gain restriction takes intermediate values, as predicted before.

Although the above example can not be taken as a demonstration of the stability restrictions in mixed delays delayed feedback amplifiers, it represents a qualitative description of the behavior of this kind of amplifiers. The complete demonstration is possible only under numerical simulation procedures. Appendix B 4 presents preliminary considerations for this demonstration.

3.4 Robustness Issues

Feedback amplifiers are affected by a wide variety of strange effects. The most immediate of these effects are perhaps, temperature, process variations (especially when dealing with integration technologies), and aging. Other kinds of specialized circuits can also experiment effects caused by environmental constraints, such as: bias noise, electromagnetic interference, radiation, pressure, or humidity, to mention some. These effects can have very different manifestations; however their effect on the amplifier's dynamics is often regarded as an expected consequence. A possible way to specify these dynamic variations is to assume that the coefficients of the characteristic equations have associated uncertainties, thus leading to the concept of uncertain systems. Fortunately, for the majority of cases, these uncertainties have bounded nature. Nevertheless, the problem of checking stability in systems with bounded uncertainties is a difficult one. The reason for this lies on the fact, that it is necessary to consider a whole family of systems instead of the nominal system alone.

Robustness is a particular area from control theory that studies the effects of these uncertainties on system's behavior, especially on stability. There are a wide variety of results concerning the robust stabilization of polynomial systems [283]. The most elegant approach was published in 1978 by Kharitonov [282, 283]. The Kharitonov's theorem establishes the necessary and sufficient conditions to achieve Hurwitz stability in systems having polynomial dynamics with constrained uncertainties on its coefficients. The major achievement of Kharitonov's theorem is to show that it is sufficient to check stability of four special Kharitonov polynomials in order to guarantee the stability of the whole set. Recent efforts extended this result to the quasi-polynomial case [144]. These results show that the number of necessary stability checks increase with the number of delays in the quasi-polynomial characteristic equation [139, 282, 283]. An important asset of the works of Silva, Datta and Bhattacharyya is that for the case of quasi-polynomial systems with only one delay the stability bounds exhibit monotonic behaviors against delay variations (see Figure 3.10, Figure 3.11 and theorems 3.6 and 3.8). Joining these results it is possible to simplify stability checks for systems with bounded uncertainties, by simply verifying stability for the largest possible delay [285].

It is possible to apply these results to the present case and thus, refining the stability bounds of delayed feedback amplifiers. However, due to the complexity and limited usage¹⁴ of such results, a different perspective over the robustness properties of such amplifiers is preferable. Such perspective can be adequately described using gain and phase margin descriptors: gain and phase

¹⁴ Limited because the principal aim of this work, is to establish conditions for maximal flat operation, that under the assumptions of theorems 3.6, 3.7 and 3.8 also imply stable operation.

margins have already been used to study both stability and robustness properties of feedback systems with one delay [140, 141, 143, 145].

3.4.1 Gain and Phase Margins

Consider the general case of a unitary delayed feedback amplifier having an open-loop transfer function specified by

$$G(s) = G_o \frac{N(-s)}{D(s)N(s)} e^{-sL} \quad (3.60)$$

Equation (3.60) captures the essence of both pure delay and all-pass delay dynamics. The gain and phase margins for such amplifier are defined respectively by Δ_G and Δ_ϕ where,

$$\begin{aligned} \Delta_G &= \frac{1}{|G(j\omega_p)|} \\ \arg[G(j\omega_p)] &= -\pi \\ \Delta_\phi &= \arg[G(j\omega_g)] + \pi \\ |G(j\omega_g)| &= 1 \end{aligned} \quad (3.61)$$

$$(3.62)$$

where, ω_p and ω_g are respectively the phase and gain crossover frequencies. The gain crossover frequency can be found by solving the auxiliary equation in (3.62), which for the present case, consists on the solution of a polynomial equation of order $2n$, $|D(j\omega_g)|^2 = G_o^2$.

The phase crossover frequency is not so easily computed. In general it is possible to apply the modified Hermite-Biehler theorem, thus finding both ω_g and $G_o(osc)$ (where $G_o(osc)$ stands for the G_o upper stability bound). When $L=0$, it is also possible to find both of these values using the Routh-Hurwitz criterion. Assuming that G_o takes the particular value $G_o(osc)$, according to the definitions of Δ_G and Δ_ϕ , the cross over frequencies must coincide, resulting in,

$$G_o(osc) = |D(j\omega_p)| \quad (3.63)$$

Equation (3.63) states that $G_o(osc)$ is solely determined by the magnitude $D(s)$ evaluated at the phase (or gain since they coincide for this situation) cross over frequency. Using (3.63) it is possible to write a meaningful form for Δ_G linking the particular value $G_o(osc)$ to any arbitrary value of G_o , indeed using (3.61), (3.60) and (3.63),

$$\Delta_G = \frac{|D(j\omega_p)|}{G_o} = \frac{G_o(osc)}{G_o} \quad (3.64)$$

Equation (3.64) allows the particularly useful interpretation of the gain margin as the ratio

between the loop-gain necessary to achieve marginal stability against some prescribed loop-gain. In this sense Δ_G shows how much G_o can be enhanced without losing stability.

3.4.2 Robustness Properties under Maximal Flat Conditions

Equation (3.64) provides an insightful perspective over the robustness characteristics of delayed feedback amplifiers under the maximal flatness constraints. Assuming that the loop-gain is adequately set to its maximal flat value, the gain margin specifies for this situation how much the loop-gain can be increased before oscillation. Furthermore, since both $G_o(osc)$ and $G_o(flat)$ (G_o under maximal flatness constraints) depend on the same dynamic parameters, they are affected by the same uncertainties. This means that, if the amplifier is appropriately tuned on a safe operating region, where both $G_o(osc)$ and $G_o(flat)$ have the similar monotonic behavior, the effect of uncertainties is reduced. Figure 3.10 and Figure 3.11 shows that this similar monotonic behavior is present on both 1st and 2nd order delayed feedback amplifiers, as the pole-delay product varies (natural frequency-delay product for the 2nd order case). For the 2nd order case, the safe operating region also depends on the damping coefficient (ξ), becoming wider for large values of ξ . For all-pass delay feedback amplifiers this statement can be formally verified, since both G_o values are known in a closed form (for 1st and 2nd order). According to Table 3.1 the gain margin is easily found.

$$\Delta_G = 4 \left(1 + \frac{T}{a_1} \right)^2 \quad (3.65)$$

for 1st order amplifiers, and,

$$\Delta_G = \frac{\xi}{G_2 \omega_n T (1 + \xi \omega_n T)} + \frac{2\xi (\omega_n T + 2\xi)}{(2\xi^2 - 1)(1 + \xi \omega_n T)} + \frac{1 + 2\omega_n T (\omega_n T + 2\xi)}{(1 + \xi \omega_n T)} \quad (3.66)$$

for 2nd order amplifiers.

Since T/a_1 is directly proportional to the pole-delay product, equation (3.65) states that the gain margin is a strictly increasing function of the pole-delay product. Nevertheless, for small values of the pole delay product (where $T/a_1 \ll 1$), the gain margin is essentially fixed, thus establishing the region for safe operation where the uncertainties have no effect on the amplifier's stability. The same behavior is also found for the 2nd order case. From equation (3.66) it is readily seen that the gain margin has three regions with distinct behavior as $\omega_n T$ is varied: on the first region, the first term in (3.66) prevails, and Δ_G decreases as the $\omega_n T$ increases; on the second region Δ_G is essentially dependent on ξ , as predicted by the middle term in (3.66); on the third region Δ_G increases with $\omega_n T$, as predicted by the third term in (3.66).

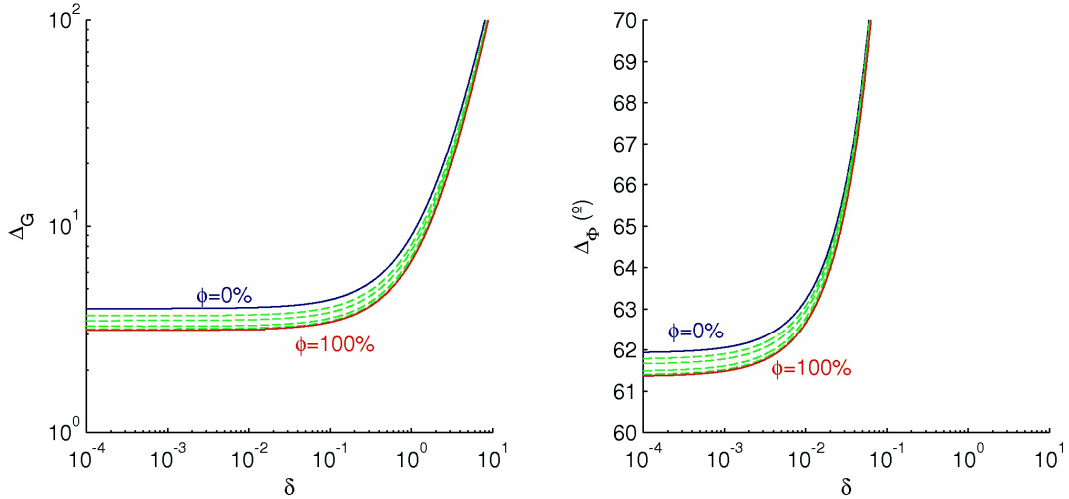
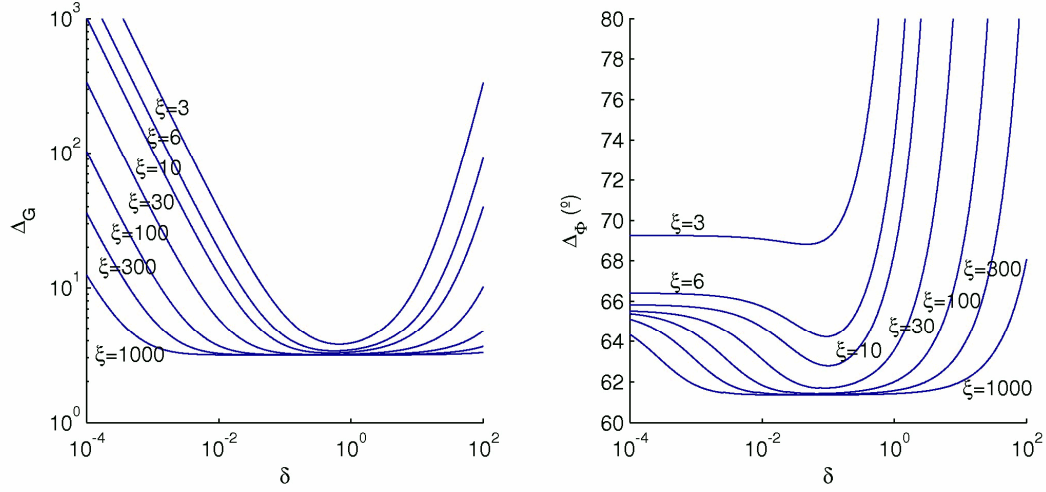
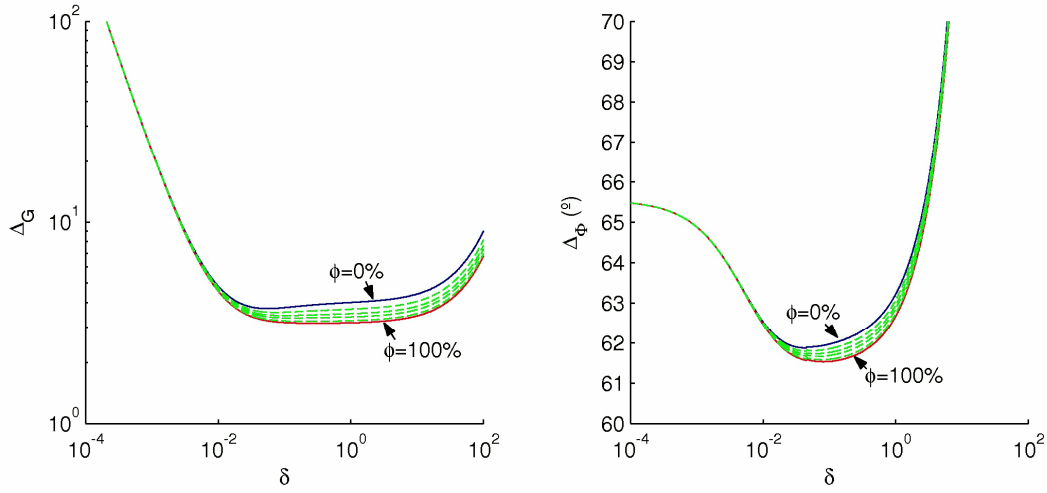


Figure 3.14 – Gain and phase margin for a 1st order delayed feedback amplifier with mixed delays.

3.4.2.1 1st Order Delayed Feedback Amplifiers

Figure 3.14 displays the gain and phase margins for a 1st order delayed feedback amplifier with mixed delays. In order to explore the nature of the delays around the feedback loop, the modeling approach of (3.59) was adopted, thus allowing to interchange the delay between pure delay and all-pass delay, having $\phi=100\%$ and $\phi=0\%$ respectively. As Figure 3.14 shows, the gain margin increases with the effective pole-delay product. This increase is also patent on Figure 3.10, and represents the fact that for increasing values of the pole-delay product, the stability bound approaches unity, while the required maximum flatness gain approaches zero. Obviously, from the circuit design perspective, negative feedback should employ large loop-gain values in order to effectively explore the feedback benefits. Thus, high gain margin values are not useful for circuit design since G_o becomes less than unity very rapidly. For small values of the pole-delay product ($\delta < 0.1$), both gain and phase margin are essentially constant: all-pass delay amplifiers ($\phi=0\%$) have a gain margin of approximately 4 and a phase margin near 62° ; pure delay amplifiers ($\phi=100\%$) have slightly less gain and phase margins, ~ 3.14 and $\sim 61.5^\circ$. Intermediate values of parameter ϕ (dashed green lines) exhibit intermediate values for both gain and phase margins, which reinforce that: mixed delays amplifiers have an intermediate nature with pure delay and all-pass delays as extreme cases; and for the same effective delay, both pure delay, all-pass delay and mixed delays amplifiers are stable under maximal flatness constraints. Common practice advises the use of gain margins between 3 and 4 and phase margins between 45° and 60° [140, 141, 143, 145, 257], leading to the final observation that maximal flat design agrees well with both of these directives.


 Figure 3.15 - Gain and phase margin for a 2nd order delayed feedback amplifier with pure delay.

 Figure 3.16 - Gain and phase margin for a 2nd order delayed feedback amplifier with mixed delays, for $\xi=50$.

3.4.2.2 Higher Order Delayed Feedback Amplifiers

Figure 3.15 depicts the gain and phase margins for the 2nd order case, assuming the presence of pure delays and several values of ξ . This figure agrees with the previous discussion relating equation (3.66). The constant gain margin range is strongly dependent on ξ , and attain a minimum value of ~ 2 . This observation shows that 2nd order delayed feedback amplifiers operating under maximal flatness have smaller gain margins. Decreasing ξ also degrades the robustness characteristics of the amplifier; the gain margin becomes strongly dependent on the pole-delay product (for 2nd order amplifiers, $\delta = \omega_n L$), which implies less immunity to uncertainties.

Figure 3.16 shows that the intermediate nature of mixed delays feedback amplifiers is also present for 2nd order amplifiers; both gain and phase margins exhibit intermediate values between the gain and phase margins of all-pass delay ($\phi=0\%$, blue line) and pure delay ($\phi=100\%$, red line)

amplifiers. This behavior was confirmed for several different values of ξ .

Higher order dynamics were also investigated leading to the general observation that higher order delayed feedback can achieve flat frequency response operation with worse robustness characteristics.

3.5 Bandwidth Optimization

There are two reasons to consider delays in the design of feedback amplifiers: stability and bandwidth. From the stability point of view, delays turn stability checks more complex. In this sense, it is easier to consider feedback systems with simple polynomial dynamics. Nevertheless, the stability bounds of polynomial systems are very optimistic, since delays due exist and may compromise stability for these cases. The other reason to consider delays is the possibility of achieving higher bandwidths using appropriate design procedures. Effectively delays in feedback amplifiers can be thought as peaking elements, allowing significant bandwidth extensions.

Conditions to achieve maximum flatness were presented on section 3.2. Following these conditions, several results were derived for 1st and 2nd order delayed feedback amplifiers on section 3.3. The remaining issues are the design procedures to meet both optimal bandwidth and flat frequency response. There are two possible procedures: delay control – when the loop gain is constrained and assuming that is possible to control the delay; and gain control – when there is a inherent difficulty to model or control the delay. Both of these procedures rely on two available features: BWER curves and flat gain conditions.

3.5.1 Bandwidth Enhancement functions

As stated in section 3.1.1, finding the cut-off frequency of delayed feedback amplifier relies on numerical procedures. Even for 1st order amplifiers the cut-off frequency results from the first solution of a transcendental equation. This makes bandwidth estimation a difficult and time consuming task. Two alternatives to abbreviate this difficulty rely on tabulated values or graphical drawing curves. Both of these methods demand a normalization scheme. On previous sections, a frequency normalization using the variable change $\alpha = \omega L$ lead to the definition of the pole-delay product as a critical parameter on the analysis. The BWER of a delayed feedback amplifier was defined in section 3.1.3, as the ratio of cut-off frequencies between the delayed feedback amplifier and the corresponding reference feedback amplifier without delay. These two points form the ground basis for the definition of a family of design curves.

In order to generalize results it is necessary to consider that the total delay around the feedback loop may have distinct origins and types. Two delay types are possible, as presented in section

3.1.1, pure delays or all-pass delays. Regarding origin, delays can be intrinsic – when originated within the amplifier elements; or extrinsic – when intentionally added. Intrinsic delays play an important role for several reasons, namely: i) establish the extreme stability restrictions; ii) establish the need for compensation or the opportunity for optimization; iii) establish the real reference amplifier - when the intrinsic delay is concentrated on the forward path between the amplifier's input and output ports, the maximum attainable BWER is less than the value predicted using the reference amplifier without delay.

Assuming that the total intrinsic delay (due to active elements, RHP zeros and circuit length) is represented by L_i , the total extrinsic delay by L_e , and the total delay around the loop by L , where $L=L_i+L_e$, the general BWER definitions for 1st and 2nd order amplifiers are,

$$BWER = \frac{\alpha_c(G_o, \delta)}{\alpha_c(G_o, \delta_i)} \quad (3.67)$$

$$BWER = \frac{\alpha_c(G_o, \xi, \delta)}{\alpha_c(G_o, \xi, \delta_i)} \quad (3.68)$$

Where $\alpha_c()$ is the normalized cut-off frequency, δ is the pole-delay product (L/a_1 for 1st order and $\omega_h L$ for 2nd order amplifiers), and δ_i is intrinsic pole-delay product (L_i/a_1 for 1st order and $\omega_h L_i$ for 2nd order amplifiers). It is possible to analyze the effects of intrinsic delays and the impact of the delay type on the BWER under maximal flat conditions, using definitions (3.67) and (3.68).

3.5.1.1 1st Order BWER Gain Curves

Figure 3.17 represents the maximum BWER for both pure and simulated delay 1st order feedback amplifiers, under maximal flat frequency response conditions. It is possible to explore a bandwidth improvement of ~ 2.25 referring to the non-delay reference situation. Using all-pass delays enables slightly less bandwidth improvement (near ~ 2.13). Figure 3.17 also shows that for both cases the BWER is almost fixed for pole-delay products less than 0.1 , which naturally agrees with the constant gain margin region for this kind of amplifiers (as shown on Figure 3.14). Pole-delay products larger than 0.1 exhibit less BWER and smaller loop-gains.

It is clear that the presence of intrinsic delays

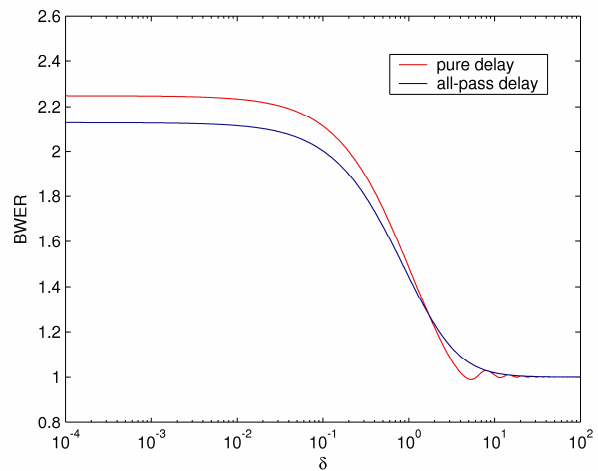


Figure 3.17 – 1st order BWER gain curves for pure and all-pass delay amplifiers.

does not affect the maximum flatness condition, as long as the predicted total delay around the loop remains larger than the intrinsic delay. However, from BWER perspective, the presence of small intrinsic delays produces larger bandwidths than the simple non-delayed reference amplifier. According to Figure 3.17, this effect is worsened as the intrinsic delay approaches the total permitted delay (both loop-gain and delay are constrained under maximum flatness conditions).

Figure 3.18 exhibits the effect of BWER reduction due to intrinsic delays, assuming that both intrinsic (L_i) and extrinsic (L_e) delays can be modeled as pure delays¹⁵, and L_i expressed as a fraction ηL of the total delay. As shown, BWER decreases down to unity as η increases towards 100%. Fortunately, the flat BWER region is preserved for all possible values of η .

3.5.1.2 2nd Order BWER Gain Curves

Following the same approach as for the 1st order case, Figure 3.19 and Figure 3.20 show the effect of adding a second pole to the amplifier dynamics. As previously, the pole-delay product is replaced for these cases by the natural frequency-delay product, and the damping coefficient is taken as a measure of the separation between poles. As depicted on Figure 3.19, for 2nd order feedback amplifiers, the BWER gain curves have a bell shaped form, meaning that for small values of δ the BWER also decreases, as the two dominant poles approach their desired maximal flat positions. As for the 1st order case, simulated delay amplifiers exhibit less BWER than pure delay

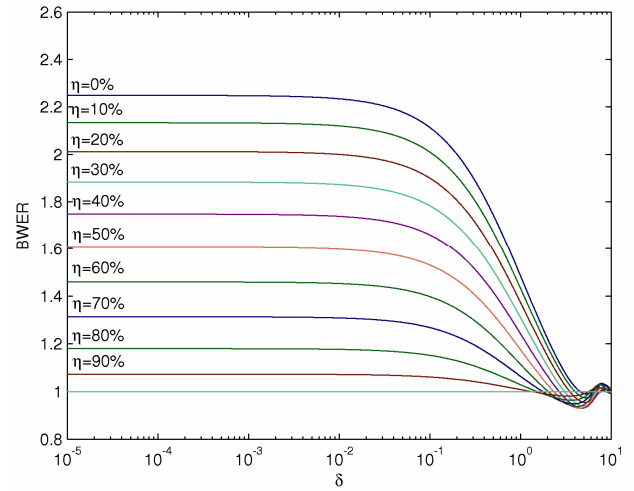


Figure 3.18 – 1st order BWER gain curves: effect of intrinsic delay.

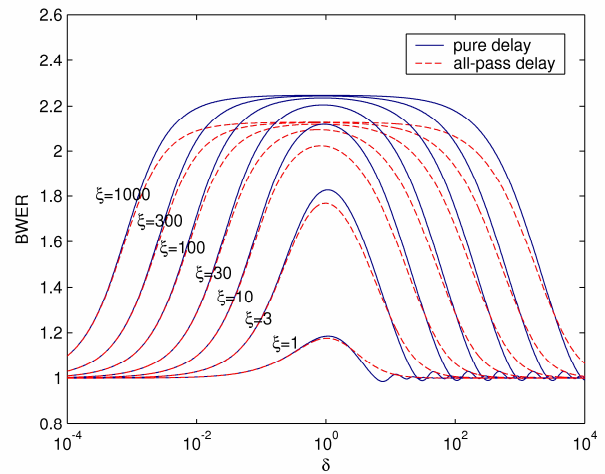


Figure 3.19 – 2nd order BWER gain curves for pure and all-pass delay amplifiers.

¹⁵ Similar results follows for all-pass or mixed delays amplifiers.

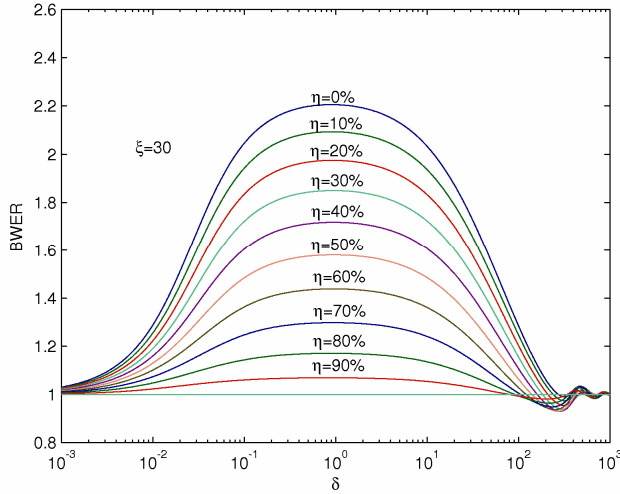


Figure 3.20 – 2nd order BWER gain curves: effect of intrinsic delay.

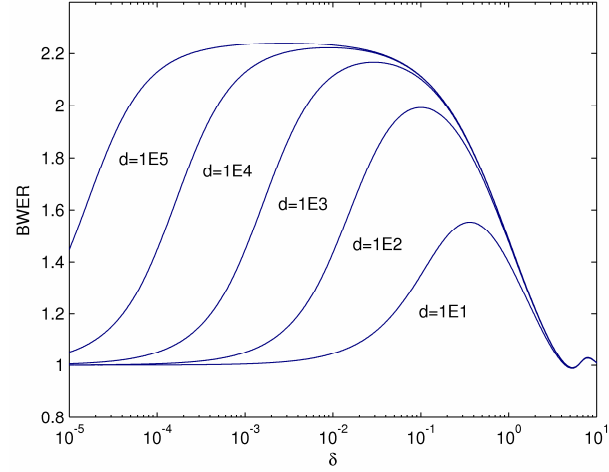


Figure 3.21 – 2nd order BWER gain curves: using the pole separation.

amplifiers. The bell shaped form exhibited by the BWER curves is also observed (in a inverted form) for the gain margin curves (see Figure 3.15). The maximum BWER occurs for similar δ intervals as the flat gain margin, meaning that the maximum BWER occurs for the more robust condition.

Figure 3.20 shows the BWER reduction effect due to intrinsic delay, in a 2nd order delayed feedback amplifier having a fixed damping coefficient of $\xi=30$ and variable pole-delay product. In 2nd order delayed feedback amplifiers these two factors – higher order dynamics and intrinsic delay act together as BWER reduction effects, thus restricting the range of applicability of this delayed feedback technique.

Since the damping factor is only an indirect measure of the pole separation, it would be more valuable from the design point of view to consider the pole separation itself. This is further supported by the fact that on the previous results, a minimum value of $\xi=1$ was often used. As it is well known, this corresponds to the situation when the two poles are both real and coincident. Large values of ξ , mean that the two poles are on the real axis. Considering that these two poles have angular frequencies of p_1 and dp_1 , with $d>1$, the natural frequency and the damping coefficient are expressed by,

$$\begin{aligned}\omega_n &= p_1 \sqrt{d} \\ \xi &= \frac{1+d}{2\sqrt{d}}\end{aligned}\tag{3.69}$$

Figure 3.21 shows the effect of these modifications in the BWER curves, when the dominant pole-delay product (or at least the low frequency pole), $\delta=p_1L$, is taken instead of ω_nL . One important observation is the fact that as the pole separation becomes larger, the BWER curve

approaches the 1st order case. This figure hints that 1st order amplifiers are from this point of view the limit case of higher order amplifiers when all the dominant poles (except those introduced by the delay) approach infinity.

A final observation shows that it is possible to predict the closed-loop frequency response using Figure 3.17 and Figure 3.21 together with the appropriate maximum flatness gain condition. These curves are also useful for design purposes, since they show the more favorable values of both BWER and necessary pole-delay product.

3.5.2 Delay Optimization

Delay oriented bandwidth optimization is the simplest method for bandwidth optimization in delayed feedback amplifiers. The delay optimization problem can be stated as: given a open-loop amplifier having low frequency gain A_o , and 1st ($a_1=1/p_1$) or 2nd (ω_n, ξ or p_1, p_2) order dynamics plus intrinsic delay (L_i), find the necessary external delay (L_e) in order to meet maximal flatness closed-loop frequency response for a closed-loop low frequency gain of H_o . The solution to this problem relies on a four step design procedure:

- 1) Using A_o and H_o , the loop-gain is given by,

$$G_o = \frac{A_o}{H_o} - 1 \quad (3.70)$$

- 2) The total delay can be found using the appropriate maximum flat gain condition. Obviously, this depends on the type of delay sought for the design. For 1st and 2nd order amplifier the maximum flat gain condition is formalized respectively as,

$$G_o = MF(\delta) \quad (3.71)$$

$$G_o = MF(\delta, \xi) = MF_\xi(\delta) \quad (3.72)$$

Where the subscript ξ means that for 2nd order amplifiers the maximum flat condition depends on both ξ and δ . Theorems 3.6, 3.7 and 3.8 off section 3.3 present maximum flat loop-gain conditions for 1st and 2nd order delayed feedback amplifiers. The present definition on (3.71) and (3.72) follow the results of these theorems, disclosing only the important design parameters. The pole delay product, δ , is obtained inverting (3.71) or (3.72). According to theorems 3.6, 3.7 and 3.8 the loop-gain conditions are quadratic forms of δ , which means that δ can have two distinct values for each value of the loop-gain, G_o . Since these two solutions have different signs, the positive one represents the sought pole-delay product.

- 3) The total delay is found using the usual definitions, $\delta=L/a_1$ and $\delta=\omega_n L$ for 1st and 2nd order

respectively. The value of η is now computed from the values of L_i and L as, $\eta=L_i/L$. There are three cases to consider: i) $\eta<1$ – it is possible to improve the bandwidth of the closed-loop amplifier using an external delay $L_e=L-L_i$; ii) $\eta=1$ - the frequency response is already under the maximal flat condition, it is not possible to improve further the amplifier's bandwidth; iii) $\eta>1$ – it is necessary to compensate¹⁶ the closed-loop frequency response, the gain and delay are above the maximum flatness condition.

- 4) The final step estimates the closed-loop bandwidth when $\eta\leq 1$. The closed-loop bandwidth for 1st and 2nd order amplifiers is given by,

$$BW = K(\delta, \eta) \omega_c(G_o, a_1) \quad (3.73)$$

$$BW = K(\delta, \xi, \eta) \omega_c(G_o, \omega_n, \xi) \quad (3.74)$$

Where $K(\)$ represents an appropriate BWER function, and $\omega_c(\)$ is the cut-off frequency of the reference non-delayed feedback amplifier. The values of $K(\)$ are calculated using the procedures of section 3.5.1. An alternative is to use tabulated or graphical forms of $K(\)$, whenever these are available. Figure 3.22 depicts an example of the graphical method to estimate BWER on a 2nd order delayed feedback amplifier; once the appropriate values of δ , ξ and η are known, the value of $K(\)$ is easily estimated from the graphical representation of the BWER. One problem with this approach is the fact that it is not feasible to produce a collection of BWER curves for every delayed feedback case. This must consist on a special set of standardized values of ξ and η .

Design examples falling between these standard values can only have inaccurate bandwidth estimations using this approach.

This simple procedure is only useful for optimization processes. The range of applicability is strongly dependent on feedback loading and non-unilateral¹⁷ effects as shall be discussed on the next chapter. The complete design of delayed feedback

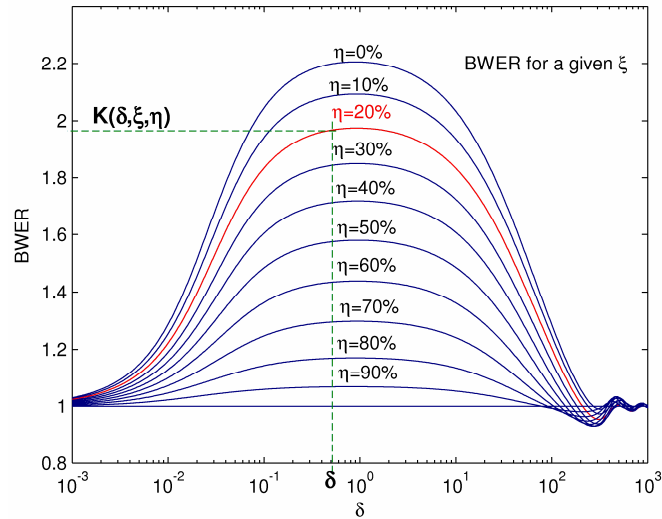


Figure 3.22 – Graphical method for bandwidth estimation.

¹⁶ Two compensation possibilities are gain reduction or dominant pole reduction.

¹⁷ Non-unilateral effects appear when some element within the feedback loop has both forward and backward transmission factors.

amplifiers demands a more evolved procedure, able to cope with these effects.

3.5.3 Gain Optimization

The problem of designing a delayed feedback amplifier can be stated as a gain-bandwidth optimization problem. The objective is to achieve a closed-loop frequency response having both maximal gain and maximal bandwidth. To achieve such goal, it is necessary to have specifications on the loop-gain dynamics: loop-gain, total delay, and poles. Since the objective includes bandwidth optimization it is possible to assume a 2nd order model and fix the pole separation factor d (or ξ instead). The general procedure should follow the next steps:

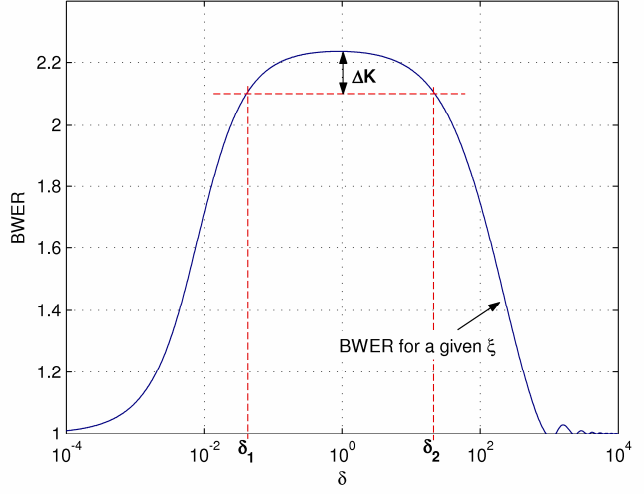


Figure 3.23 – Graphical method to find δ in a 2nd order delayed feedback amplifier.

- 1) Assuming $d > 1$ and $\varepsilon > 0$ (defined as the tolerance on BWER), it is possible to find a set of pole-delay values satisfying the following criterion (see Figure 3.23),

$$\delta_1 \leq \delta \leq \delta_2 \Rightarrow \Delta K \leq \varepsilon \quad (3.75)$$

Where ΔK represents the tolerated variation on BWER around its maximum value

- 2) Using the appropriate maximum flat gain condition (selected according to the type of delay considered on the design), the loop-gain is restricted by,

$$MF(\delta_2, \xi) \leq G_o \leq MF(\delta_1, \xi) \quad (3.76)$$

Since the objective is to maximize the closed-loop gain, the obvious choice falls on $MF(\delta_2, \xi)$, since small loop-gains correspond to higher closed-loop gains. This choice is also the best from the delay perspective, since high pole-delay products convert into high delay values, which are less restrictive under the presence of intrinsic delays. However, slightly larger values of G_o should be used in order to permit some allowance for final adjustments. It is also important to select G_o based on gain margin criteria; to this end the value of ε should also reflect that the gain margin remains essentially fixed.

- 3) Combine these two criteria with the design of the amplifier. Once the two dominant poles are known, the total delay can be estimated within the range,

$$\frac{\delta_1}{p_1} \leq L \leq \frac{\delta_2}{p_1} \quad (3.77)$$

Where, p_1 is the lowest frequency pole (in rad/s).

- 4) Using an estimate for the intrinsic delay (this can be achieved using the delay specification of the active devices and adding to this the effect of possible negative zeros), find the value of $\eta = L_i/L$ (assuming that L_i is concentrated on the forward path). Three possibilities may occur: i) $\eta > 1$ - it is impossible to attain the specifications of 1) and 2). The procedure returns to step 3); ii) $\eta = 1$ – all the specifications are met; iii) $\eta < 1$ – in order to fulfill all the specifications, it is necessary to add an external delay with value $L_e = L - L_i$.
- 5) The closed-loop bandwidth can be evaluated using (3.74).
- 6) The final step of the procedure uses simulated results (or test results if available) to fine tune the design: obtain better estimates for G_o , p_1 , d , L , L_i , and adjust the design following the above procedure.

3.6 Concluding Remarks

This chapter addressed the concept of delayed feedback. As it was demonstrated, the presence of small delays inside the feedback loop of a feedback amplifier can be used for bandwidth optimization purposes. It was demonstrated that the presence of these delays can cause frequency peaking phenomena, in a similar fashion as the methods of inductive or capacitive peaking discussed in chapter 2. The exploitation of these peaking effects could lead to significant bandwidth improvements. It was shown that a maximum bandwidth improvement of 241.4% could be attained solely by adding a delay to the feedback loop of a feedback amplifier. As for the cases of inductive and capacitive peaking, it was demonstrated how the maximal flat frequency response condition could be attained. Lemma 3.1 and theorem 3.1 establish necessary and sufficient conditions to achieve maximum flatness in delayed feedback amplifiers. As it became clear from these results, the maximal flat argument can only be maintained on a restricted frequency range. The reason for this arises due to the quasi-polynomial dynamics of delayed feedback amplifiers.

Stability and robustness issues were also addressed in this chapter. The main result showed that maximal flat gain conditions led to stable and robust closed-loop amplifiers, able to deliver more bandwidth than their non-delayed equivalents.

Finally, two algorithmic approaches able to address bandwidth optimization or gain- and bandwidth optimization problems in delayed feedback amplifiers were presented. These methods require precise modeling information about the open-loop amplifier dynamics, which will be discussed on the next chapter.

4 DELAYED FEEDBACK AMPLIFIERS – DESIGN CONSIDERATIONS

“She said: What is history? And he said: History is an angel being blown backwards into the future. He said: History is a pile of debris, And the angel wants to go back and fix things To repair the things that have been broken. But there is a storm blowing from Paradise, And the storm keeps blowing the angel backwards into the future. And this storm, this storm is called Progress” – Laurie Anderson.

Summary

Traditional feedback amplifier design techniques suffer from several inaccuracies. As exposed on the previous chapter, the presence of delays in these amplifiers can originate a whole set of deviant behaviors, that range from simple and negligible design inaccuracies to harmful oscillations. Using a maximal flat frequency response criterion leads to robust design strategies that are not only desirable but also benefic. This chapter discusses analysis and synthesis techniques suitable for the design of delayed feedback amplifiers. The aim of these techniques is specially focused on the assessment of the so called “return ratio” and on the design of simple yet effective delay elements.

4.1 General Discussion

The previous chapter presented two procedures suitable to the design of delayed feedback amplifiers. The first of these procedures is a simple bandwidth optimization scheme. Once the open-loop amplifier dynamics are completely specified, the optimization is accomplished using an optimized loop delay. The second method goes one step further. The open-loop dynamics are not known in advance, the design goals aim for both maximal gain and bandwidth. Both methods use maximal flat frequency response criteria and require a precise modeling of the open-loop amplifier dynamics.

Finding the open-loop dynamics of a feedback amplifier can be a complicated task. Considering the simple feedback model of the previous chapter, comprising a forward amplifier and a feedback path with transmission β , finding the open-loop transfer function consists in opening the feedback loop at the input of the amplifier and measuring the gain from the input to the output of the amplifier and back again to the input through the opened feedback path. However, some difficulties arise due to the simplicity of the considered model. Namely,

- 1) it assumes that both the forward amplifier and the feedback factor are ideal unilateral blocks;
- 2) it neglects the loading effects imposed by the feedback network on the forward amplifier;
- 3) it is based on the identification of two main blocks, the forward amplifier and the feedback network.

In general, real feedback amplifiers violate these three assumptions:

- 1) Feedback amplifiers are composed of one or more transistorized gain stages. Since transistors are not unilateral devices, the unilateral assumption represents an unrealistic argument on the forward path. Also, the feedback network is generally composed by passive elements, with both forward and reverse transmission factors;
- 2) The feedback network acts directly on the input and output of the amplifier. Depending on the values of both input and output resistances of the amplifier the loading effects may represent important detrimental factors;
- 3) Finally, it is often difficult to identify properly the forward gain and the feedback network in real feedback amplifiers (one such case arises in feedback amplifiers comprising unbalanced differential amplifiers on the forward path), especially if the feedback is only local (only a portion of the feedback signal is feed directly to the input. Global feedback applies directly the feedback signal to the input, such as in shunt-shunt configurations). For some of these situations, it is complex to describe the forward amplifier as a two-port network, since the reference node cannot be associated with any of the two input terminals; an alternative

representation should consider a three-port network having three input terminals (with the reference node as the common terminal). However, with such three-port description it is at least non-trivial to find a suitable feedback configuration comprising a forward amplifier and a feedback network.

There are several methods suitable to the analysis of feedback amplifiers incorporating non-unilateral blocks [75, 78, 79, 81-85, 87, 88, 92, 93, 95, 97-100, 103-105, 110, 248, 267, 277, 280]. Since their invention in 1937 by Black [75], feedback amplifiers have been analyzed by several methods, namely:

- 1) Using the concept of return ratio introduced by Bode [248]. The return ratio is a general circuit analysis tool that measures the effect on circuit performance caused by one particular element. This is a useful tool to evaluate the effects of the feedback network on the overall amplifier;
- 2) Using signal flow graphs and Mason's gain formula [79, 81, 82, 103]. In Mason's own words, "*A way to enhance writing gain at a glance*", but, this method requires an extra representation step: the original circuit has to be converted into an adequate signal flow graph and only then it will be possible to use signal flow graphs. The conversion step is somewhat obscure, since the signal flow graph would be based on several branch equations of the original circuit. A simplified conversion method was reported in [103], but even with such an optimized procedure, signal flow graphs hide important aspects from circuit design;
- 3) Using the asymptotic gain model introduced by Rosenstark [88] and modified by Choma [92, 98, 99]. This method combines return ratio and signal flow graphs techniques. It represents an efficient analysis procedure for feedback amplifiers, since it requires the measurement of only three parameters: return ratio, forward gain, and asymptotic gain (or the null return ratio using Choma modification), and still provides an exact algorithmic procedure;
- 4) Using general two-port representation. This is the preferred textbook design method [252, 264]. The widespread usage of this method arises due to the fact that the feedback amplifier is modeled as the traditional two block diagram comprising one forward amplifier and the feedback network, both described by an appropriate two-port parameter matrix (when possible; for several cases the analysis is only approximated). The major drawback of this method is the necessity of requiring a two-port description of its constituents; and finally;
- 5) Using Kirchhoff's laws applied to the analysis of the complete feedback amplifier.

One extra complication arises when it is necessary to consider the presence of delays on the overall analysis. Despite of the delay nature, it is desirable to have similar methods both for analysis and design. The results reported in chapter 3 assume that the delays can have two distinct natures:

intrinsic and extrinsic.

Intrinsic delays are due to amplifier components and physical apparatus, while extrinsic delays are intentionally placed on the amplifier for optimization purposes. Intrinsic delays are in large measure uncontrolled. Their presence arises mainly due to the transistors that composed the amplifier; it is a designer choice to include more or less transistors in his design. This choice is ultimately the best form of controlling intrinsic delays. These delays have different manifestations for different transistors types as reported in appendix A.1. Nevertheless, two physical reasons can usually explain the existence of delays in transistors: the first reason is associated to the physical phenomena of charge conduction within the transistor [157-173]; a second reason is due to the presence of inter-terminal parasitic capacitances, which for certain amplifier configurations (mainly CX stages), can originate RHP zeros (in view of the results discussed on chapter 3, RHP zeros can be considered sources of delay).

On the other hand, extrinsic delays can be designed using a multitude of available methods, for instance: using matched transmission lines with adequate length; exploring the presence of RHP zeros in circuits; using artificial transmission lines; and many more.

This chapter discusses methods suitable for analysis and design of delayed feedback amplifiers. Section 4.2 presents the asymptotic gain model. Among all the existing methods to the analysis of feedback in amplifiers, the asymptotic gain model represents the best choice, due to its general formulation, algorithmic nature and its exact results. Section 4.3 discusses the nature of intrinsic delays in transistors. Both bipolar and field effect transistors are considered in this discussion. Section 4.4 addresses the problem of delay design. Two methods for the design of delay elements are discussed: one based on meander transmission lines and other based all-pass transfer function synthesis methods. Section 4.5 discusses the theoretical aspects of the design of a transimpedance amplifier employing delayed feedback concepts. Finally, section 4.6 concludes the chapter.

4.2 Asymptotic Gain Model

The asymptotic gain model was originally introduced by Rosenstark [88]. Its origins are related with general signal flow graph theories [79, 81, 82]. The formulation of the asymptotic gain model for feedback amplifiers consist on the definition of three basic entities: 1) the return ratio – a measure of the output's effect on the input circuitry, that is, a measure of the loop-gain; 2) the asymptotic gain – the ultimate gain if the feedback is effective; 3) the direct gain – due to device imperfections, every feedback amplifier has its own direct terms outside the feedback loop.

4.2.1 Preliminary Definitions

Any generic N terminal network comprising independent current or voltage sources, passive admittances (due to resistances R , inductances L , and capacitances C), plus voltage controlled current sources (VCCS), can be defined in terms of its nodal equations¹⁸. Considering the N terminal network of Figure 4.1, this can be put in matrix form as,

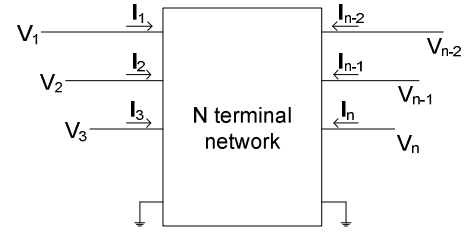


Figure 4.1 – General N port network.

$$\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} \quad (4.1)$$

Using matrix notation it can also be written $[Y][V]=[I]$, where $[Y]$ represents the nxn nodal admittance matrix. It is possible to express any nodal voltage of the complete circuit recurring to the definitions of driving point impedance and transfer impedance. The driving point impedance is defined as the impedance seen at node i when all the independent sources connected to all the other nodes of the circuit are set to zero (this means substituting all voltage sources by short-circuits and current sources by open-circuits). Then, if the voltage V_i is adequately written as function of I_i , using Cramer's rule [254] results in,

$$Z_{ii} = \frac{V_i}{I_i} = \frac{\Delta_{ii}}{\Delta} \quad (4.2)$$

where Δ is the determinant of $[Y]$ and Δ_{ii} is the cofactor¹⁹ obtained from Δ after deletion of row i and column i . Similarly, the transfer impedance between node i and j , is defined as,

$$Z_{ij} = \frac{V_j}{I_i} = \frac{\Delta_{ij}}{\Delta} \quad (4.3)$$

taking, again all the independent sources connected to other nodes except node i equal to zero. It is possible to express any relation between any pair of currents or voltages of the original network, using these definitions.

The circuit determinant Δ can be expressed as an explicit function of some element in the

¹⁸ Or mesh equations, considering current controlled voltage sources (CCVS), instead of VCCS

¹⁹ The general definition of cofactor is $\Delta_{ij} = (-1)^{i+j} \det[Y_{ij}]$, where $[Y_{ij}]$ is the minor of $[Y]$ obtained after the deletion of row i and column j .

original N-network. Taking, for instance, γ as the transconductance of a generic transistor with its X, Y and Z terminals connected to ground, node u and node v respectively as in Figure 4.2. The N-network determinant can be expanded into its minors, using Laplace expansion procedure [87, 89, 101, 294]. Since γ appears only in column u and row v of the original $[Y]$ matrix, the Laplace expansion shows that,

$$\Delta = \Delta^0 + \gamma \Delta_{vu} \quad (4.4)$$

where, Δ^0 represents the value of Δ when γ is set to zero.

Equation (4.4) is useful for the inspection of the effect of one particular element on any possible transfer function defined through equations (4.2) and (4.3). Clearly, γ may be any circuit parameter rather than the transconductance, but this will demand a different and more general formulation recurring to modified nodal analysis. Generally, transistors are modeled as VCCS (Voltage Controlled Current Source), thus justifying the above simpler model analysis.

4.2.2 Return Ratio and Return Difference

Classical results on feedback amplifiers comprise the simple β -A model (a feedback loop comprising a forward transfer function A and a backward transfer function β), where the negative of the product βA , the loop-gain ($-\beta A$), plays an essential role. It is often difficult to have a clear distinction between blocks A and β in real feedback amplifiers. As previously discussed, the necessity of a two-port representation for both β and A blocks is the main factor that originate this difficulty. Given the importance of the loop-gain in the dynamics of feedback amplifiers, it is necessary to have a clear understanding of the effect of feedback in amplifiers. One such measure that provides this notion, without relying on an explicit representation of β , is the return ratio. The return ratio represents a circuit's perspective similar to the idealized loop-gain. As the loop-gain, it is a measure without unity representing the amplifier's internal dynamics, with the advantage of having a general definition.

Considering the general set-up of Figure 4.2 and assumptions leading to (4.4), the return ratio, R , for any given element γ inside the original N-network is defined by the product of its transmission γ (between terminals u and v) by its backward transfer with γ set to zero. Using matrix notation [248], this is

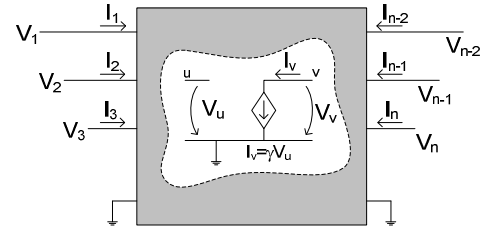


Figure 4.2 – N-Network with element γ between nodes u , v and reference.

$$R = \gamma \frac{V_u}{I_v} = \gamma \frac{\Delta_{vu}}{\Delta^0} \quad (4.5)$$

Clearly, the cofactor Δ_{uv} has no dependence on γ . Another important quantity is the return difference or feedback, defined as $F=I+R$, given by,

$$F = 1 + \gamma \frac{\Delta_{vu}}{\Delta^0} = \frac{\Delta^0 + \gamma \Delta_{vu}}{\Delta^0} = \frac{\Delta}{\Delta^0} \quad (4.6)$$

Equation (4.6) states that the return difference is just the ratio between the nodal admittance determinants of the complete circuit, against the circuit having the element effect set to zero. The importance of this result is that in order to measure the return difference, or the return ratio, it is possible to choose any convenient element γ inside the circuit, without having to clearly identify the amplifier and the feedback circuitry.

For the simple classical model of Figure 4.3, the return difference is a measure of the voltage returning to the input, that is $F=X_o/X_i=1-\beta A$. For this situation the return ratio and the loop-gain

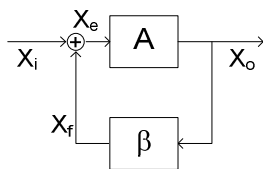


Figure 4.3 – Classical β -A model.

have the same physical significance, since $R=-\beta A$. However, it should be clear that for real amplifiers the return ratio results different from the classical loop-gain due to the factors discussed in section 4.1. The assumption of negative feedback implies that there must be an odd number of phase reversals inside the feedback loop, thus R represents a positive value (the same observation is also true relating to the model of Figure 4.3 and the quantity $-\beta A$).

4.2.3 Asymptotic Gain Model Formulation

Rosenstark based his asymptotic gain model on the return ratio definition [88]. Its derivation is

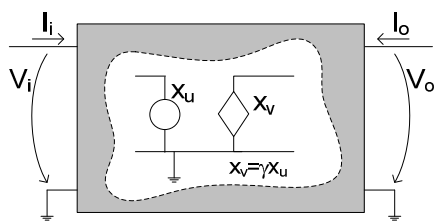


Figure 4.4 – General feedback amplifier.

completely general and can be preformed using simple circuit analysis or signal flow graph theory. Figure 4.4 represents a general feedback amplifier, with its input variables V_i and I_i , output variables V_o and I_o , and an internal controlled source of any kind. Assuming that X_I represents one of the input variables $\{V_i, I_i\}$, X_n represents one of the output variables $\{V_o, I_o\}$ and x_u, x_v represent the control and controlled voltage (or

current) pair. Choosing X_I and x_v as independent sources, it is possible to write the following matrix equation,

$$\begin{bmatrix} Y_n \\ x_u \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_l \\ x_v \end{bmatrix} \quad (4.7)$$

Assuming that the desired transfer function H , relates Y_n against X_l , and that $x_v = \gamma x_u$, results in,

$$H = \frac{Y_n}{X_l} = \frac{-\gamma A_{22}}{1 - \gamma A_{22}} \left(A_{11} - \frac{A_{12} A_{21}}{A_{22}} \right) + \frac{A_{11}}{1 - \gamma A_{22}} \quad (4.8)$$

The significance of each factor in (4.8) follows directly by inspection [88]. The asymptotic gain, G_∞ , is defined as the limit of H when γA_{22} becomes very large, that is,

$$G_\infty = \lim_{\gamma A_{22} \rightarrow \infty} H = A_{11} - \frac{A_{12} A_{21}}{A_{22}} \quad (4.9)$$

The direct gain term, G_0 , is defined as the remaining gain when γA_{22} approaches zero, that is,

$$G_0 = \lim_{\gamma \rightarrow 0} H = A_{11} \quad (4.10)$$

Finally, setting the input variable to zero ($X_l=0$) and replacing x_b by the value γ , the second equation in (4.7) gives $x_a = \gamma A_{22}$, which means that, the term $-\gamma A_{22}$ is by definition [88, 248] the return ratio of the feedback amplifier. Joining all the above considerations, the final asymptotic gain formula for feedback amplifiers is given by,

$$H = G_\infty \frac{R}{1 + R} + \frac{G_0}{1 + R} \quad (4.11)$$

Equation (4.11) has all the important parameters involved in the design of a feedback amplifier: G_∞ quantifies the final gain if the feedback is effective (R larger than 1), R specifies the amplifier dynamics and G_0 represents a non-ideal term due to non-unilateral factors within the amplifier.

The asymptotic gain model provides an exact and algorithmic method to analyze feedback amplifiers. From a analytic point of view the parameters in (4.11) are easily accessed using the following three step procedure: 1) set the control parameter γ to zero, the direct gain term G_0 is the desired transfer function between input and output under this condition; 2) the return ratio can be evaluated using its definition, or recurring to the methods summarized on appendix C 1; finally, 3) the asymptotic gain term is accessed setting the control parameter to infinity and calculating the transfer between input and output.

This three step procedure is also useful for measuring and modeling purposes. The first two steps are easily implemented using procedures similar to the ones used to measure the return ratio. However, the third may pose a challenging problem: in real amplifiers it is difficult to make the control parameter infinite. Nevertheless, it is possible to have a different interpretation of this condition; using $x_v = \gamma x_u$ and solving (4.7) in order to x_u , gives,

$$x_u = \frac{A_{21}}{1 - \gamma A_{22}} X_1 \quad (4.12)$$

Setting γ to infinity is equivalent to setting the control variable x_u to zero for finite values of the input X_1 (this is reinforced by fact that the controlled variable x_v must be finite). Killing the control

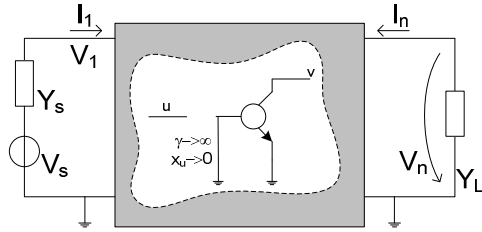


Figure 4.5 – G_∞ measurement.

variable must be realized with special care. Since x_u may not be an independent source (in general it is not), killing x_u is not simply accomplished with an open or short circuit on the specified node (or branch). As before, if the control parameter is the transconductance parameter of one transistor within the amplifier, killing x_u is equivalent to

shorting the Y terminal of the transistor and removing all its influence from node u (x_u is, for this case, the voltage measured on the transistor's input resistance r_i , thus killing x_u is equal to preventing any current to flow on this resistance). Figure 4.5 depicts the set-up to measure G_∞ ; comparing with the original amplifier set-up of Figure 4.2, the influence of γ and all the transistor input loading on node u has been effectively removed.

The procedure to measure G_0 is similar to the set-up of Figure 4.5, with the difference that setting γ to zero does not remove the transistor's influence on node u , as depicted. The same considerations applied on the measurement of R are also needed for this case.

4.3 Delays in Active Devices

The existence of time-delays in active electronic devices is a natural consequence of the current conduction phenomena. As in their passive counterparts, current is defined as the amount of charge flowing per unit of time. Since electrons cannot move with unlimited velocities, there is always a time to travel from one point to the next within the active device. Unfortunately the complete description of delay phenomena within active devices, more precisely, within transistors requires more detail than this.

The principal reason for this increased complexity lies in the fact that current conduction in semiconductors is not simply due to electrons; their absence is another possible form of charge transfer. This is conveniently described using a simple model with two types of charge carriers: electrons and holes (the absence of an electron). There are also two types of semiconductor materials: N type – having an excess of free electrons (electrons that are not strongly bounded to the semiconductor structure and thus, can acquire sufficient external energy to move freely); and, P type – having an excess of holes (representing an excess of positive charge and a trap where electrons may fall giving raise to the recombination process). N type and P type semiconductor

materials are obtained from a sample of intrinsic semiconductor crystal through a process named doping. Doping consist on the addition of impurity atoms to the intrinsic semiconductor, thus forming an extrinsic semiconductor. If these impurity atoms have one extra valence electron than the atoms of the intrinsic semiconductor under consideration, the composed material will exhibit an excess of free electrons (N type semiconductor). If, on the other hand, the impurity atoms have one less valence electron than the atoms of the intrinsic semiconductor, the composed material will exhibit an excess of holes (P type semiconductor).

Since electrons are the ultimate charge carriers, conduction in type P semiconductors is performed on a recombination-generation basis: forcing an electron current into a type P semiconductor bar, originates a series of recombination-generation processes, where the free electrons occupy the vacancies on the crystal structure – recombination; and trapped electrons acquire sufficient energy to escape from crystal bounds to form a electron-hole pair – generation. The recombination-generation process is quite different from current conduction in conductors. One relevant difference is the fact that it requires more time to achieve the same conduction effect. In fact electrons in semiconductors have smaller velocities than electrons in conductor materials, and this is even worse for P type semiconductors. The measure of the ability of an electron to move within a semiconductor submitted to an externally applied electric field is called its mobility; the mobility and the electric field define the electron velocity on both types of materials.

Semiconductor currents induced by externally applied electric fields are known as drift currents. The magnitude of these currents is proportional to the mobility of carriers (electrons or holes), the electric field strength, the electronic charge and the carrier density. Due to the mechanism of recombination-generation and to the fact that electrons tend to occupy lower energy levels, is possible to have another current conduction mechanism. The diffusion currents occur whenever there is a charge density gradient within the semiconductor. The charge carriers tend to move from places where they exist in large concentrations to places where they are in minority. These two mechanisms contribute to define the average velocity of charge carriers inside semiconductor.

Presently, several semiconductor materials are used to fabricate transistors. Silicon (Si) and germanium (Ge) were the first materials used as semiconductors. Nowadays, other kinds of compounds are used to form semiconductors, such as: gallium-arsenide (GaAs, used mainly in MESFET²⁰ and HEMT²¹ transistors), silicon-germanium (SiGe, permits direct integration within standard CMOS (Complementary Metal-Oxide-Semiconductor) process, furnishing these processes

²⁰ MESFET – MEtal Semiconductor Field Effect Transistor.

²¹ HEMT – High Electron Mobility Transistor.

with high frequency HBT²² transistors), indium-phosphate (InP used for very high frequency applications, in both HEMT, P-HEMT²³ and HBT transistors) and gallium-nitrate (GaN, used in power demanding applications and high efficiency LEDs). One immediate behavior distinction between compound semiconductors and the simple silicon or germanium is the Gunn effect. This effect traduces in the presence of negative derivative on mobility versus electric field dependency [274], and it is also one of the major advantages of these semiconductors. Simple semiconductors exhibit velocity saturation: the proportionality between the mobility and the electric field degrades, as the electric field becomes more intense. The same effect is also observed in compound semiconductors for intense applied electric fields. However, for moderate electric fields, due to Gunn effect, it is possible to explore larger carrier velocities than those verified in simple semiconductors. The higher velocity promise of compound semiconductors makes these type of materials specially suited for the fabrication of very high frequency transistors and other devices.

The above discussion presented a schematic overview of some semiconductor effects that have paramount impact on device propagation characteristics, and hence on device delays. A wider discussion of these effects is unfortunately outside the scope of the present work. More organized and comprehensive information relating these topics can be found in [253, 260, 262, 272, 274, 275]. The following analysis is focused on common transistor's available delay models, divided into: bipolar transistors [158-160, 165, 167, 170, 171, 173] and field effect transistors [163, 164, 166, 169].

4.3.1 Delays in Bipolar Transistors

Figure 4.6 depicts a vertical bipolar transistor, commonly found in integrated circuits. The bipolar transistor is a three terminal device consisting of a collector, a base and a emitter. There are

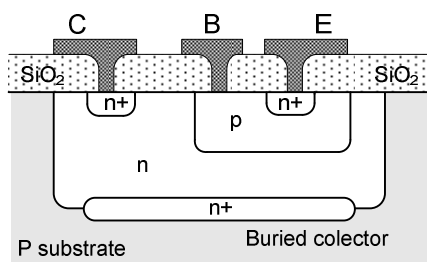


Figure 4.6 - Vertical bipolar transistor.

two types of transistors according to the arrangement of semiconductor materials: NPN transistors (similar to the structure in Figure 4.6), comprising three layers of alternating semiconductors where the collector and emitter are connected to type N semiconductor and the base is connected to the middle type P layer; and PNP transistors having the

same arrangement as before but using type P semiconductors for the base region.

²² HBT – Heterojunction Bipolar Transistor

²³ P-HEMT – Pseudomorphic HEMT

When two semiconductors of different types are joined together, a P-N junction appears. Bipolar transistors comprise two such semiconductor junctions, between base-collector (BC) and base-emitter (BE). These two junctions form the basis of the bipolar transistor's operation. At equilibrium, with no external voltages applied, the junctions are depleted of free carriers. Since N type and P type have excess of electrons and holes, respectively, these free carriers recombine near the junctions, forming two depletion layers. These depletion layers give rise to two potential barriers, having positive charge on one side and negative on the other. The built in potential of these barriers restrains the recombination process, since the free electrons and holes on both sides of the P-N junctions have to acquire more and more energy to cross these barriers and recombine. Under this no voltage condition, no current can flow between the three terminals. Applying two external voltage sources, one with the same direction as the BE junction (positive for NPN transistors and negative to PNP) and other with opposite direction as the BC junction, results in two different phenomena: i) the external voltage V_{BE} (applied to BE junction) narrows the BE potential barrier favoring the charge exchange between the emitter and base regions, resulting in a excess of minority carriers (electrons on P type and holes on N type materials) coming from the emitter into the base region. These minority carriers diffuse across the base region and reach the BC junction; ii) the V_{BC} voltage widens the potential barrier of the BC junction, thus preventing any exchange of minority carriers into the base region. However, the minority carriers traveling from the emitter into the base region and reaching the BC potential barrier, are swept across the BC junction into the collector region. Some of these minority carriers recombine in the base region; however the major portion reaches the collector region forming the collector current. Transistors operating in this condition are said to operate in the forward active region. Of course, the same conduction mechanism is possible with BC directed biased and BE reverse biased (reverse active region), but due to device geometry, this is less efficient.

Following the above discussion, the important aspect to retain is that current conduction in transistors is mainly due to the diffusion of minority carriers in the base region. It is possible though, to have a joint effect of drift and diffusion currents. The above reasoning is true under uniform doping levels of the base semiconductor. Using an appropriate doping profile the electric field on the base region can be adequately controlled [158-160]. The major benefit of having a drift current component is to decrease the base transit time and hence enhance the high frequency capabilities of the device. Under a constant electric field, the current gain between collector and emitter in a common base configuration is approximately given by,

$$\alpha(j\omega) = \frac{Ze^\eta}{\eta \sinh Z + Z \cosh Z} \quad (4.13)$$

where,

$$\eta = \frac{1}{2} \frac{E}{V_t} W_B \quad (4.14)$$

and,

$$Z = \sqrt{\eta^2 + \frac{W_B^2}{L_D^2} (1 + j\omega\tau)} \quad (4.15)$$

with V_t representing the thermal gap voltage, W_B the base width, E the electric field, L_D the diffusion length of the minority carriers in the base region, τ the minority carrier lifetime and ω the angular frequency. Equation (4.13) results from considering the base region a distributed transmission medium [160]. Figure 4.7 represents the normalized current gain in a common base configuration, assuming η as doping dependent parameter, and $\omega_o = L_D^2 / W_B^2 \tau$. As can be seen the existence of an electric field within the base region favors the high frequency performance of the transistor. However, this also adds an excess phase term to the current gain.

Usually, there is no need to consider the distributed phenomena within the base region of a transistor. For these situations equation (4.13) represents a needless complication factor in circuit analysis. Simpler small signal models, like the Π or T models [158], are sufficiently accurate for applications ranging to moderate frequencies of operation (see appendix A 1). In these models the frequency dependent parameters are modeled appending linear capacitances between the base terminal and the emitter and collector terminals. Using the T-model, assuming a fixed current gain α_o , the short-circuit current transfer between emitter and collector with base grounded is given approximately by,

$$\alpha(j\omega) = \frac{\alpha_o}{1 + j\omega r_e c_\pi} \quad (4.16)$$

where $r_e = \alpha_o / g_m$, and $\omega_\alpha = 1 / r_e c_\pi$ is the α cut-off frequency²⁴. Figure 4.8 presents a comparison between (4.13) and (4.16). The magnitudes of (4.13) and (4.16) are very close for frequencies one decade above ω_α . However this simple circuit level approximation can not predict accurately the

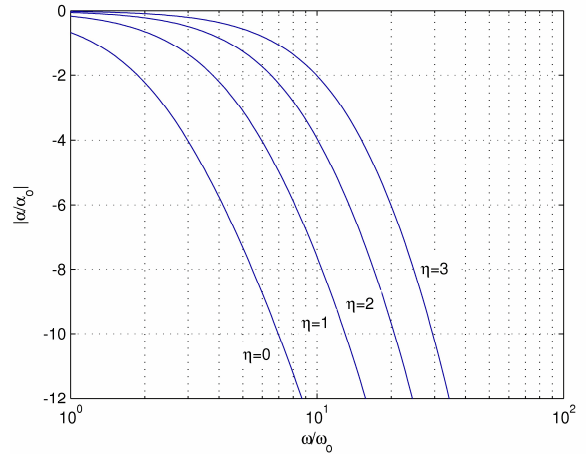


Figure 4.7 – Normalized common base current gain for different base doping profiles.

²⁴ The α cut-off frequency is also approximately related to the frequency normalization factor ω_b , using $\omega_\alpha \approx 2.4 \omega_b$.

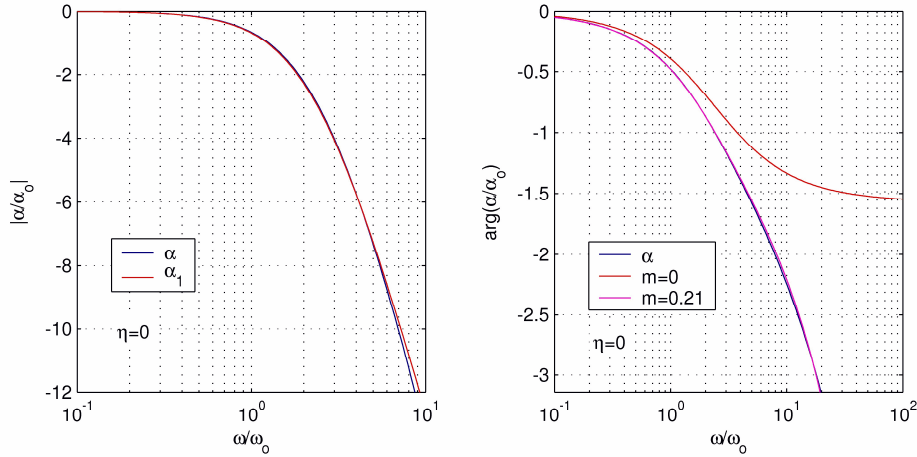


Figure 4.8 – Normalized common base short-circuit current gain, complete and approximated models for $\eta=0$; magnitude and phase.

phase variation, since it only predicts a total phase variation of 90° . In [158-160] it was derived a sufficiently simple approximation of (4.13) using implicitly the circuit level result of (4.16), which can be written as,

$$\alpha(j\omega) = \frac{\alpha_o e^{-jm\omega/\omega_\alpha}}{1 + j\omega/\omega_\alpha} \quad (4.17)$$

where m is a factor depending on the parameter η . As can be seen in Figure 4.8, the phase variation for $m=0$ (predicted from (4.16)) levels off at -90° ($-\pi/2$ radians), while the real phase variation continues to decrease with frequency. Using $m=0.21$ in (4.17) the phase variation of (4.13) is accurately approximated (at least for frequencies extending one decade above ω_α). It is straightforward to show that the excess phase term represented by the complex exponential term in (4.17) also exerts its influence on the common emitter short-circuit current gain. In fact,

$$\begin{aligned} \beta(j\omega) &= \frac{\alpha(j\omega)}{1 - \alpha(j\omega)} \\ &\approx \beta_o \frac{e^{-jm\frac{\omega}{\omega_\alpha}}}{1 + j\frac{\omega}{\omega_\alpha} \frac{1 + m\alpha_o}{1 - \alpha_o}} \end{aligned} \quad (4.18)$$

where $\beta_o = g_m r_\pi = \alpha_o / (1 - \alpha_o)$, and the exponential term on the denominator was approximated by a first order Taylor expansion [254]. The resulting β cut-off frequency is related to the Π model parameters by

$$\frac{\omega_\beta}{\omega_\alpha} = \frac{1 - \alpha_o}{1 + m\alpha_o} = \frac{1}{1 + \beta_o} \left(1 + \frac{c_\mu}{c_\pi} \right) \quad (4.19)$$

In conclusion, the phase variation of a real transistor can not be accurately described by the

simple small signal Π and T models without an extra delay term associated with: the current gain of the CCCS (Current Controlled Current Source) of the T model, or the transconductance gain of the VCCS (Voltage Controlled Current Source) of the Π model.

HBT transistors (see Figure 4.9) have slightly more complicated delay models, due to the intrinsic properties of compound semiconductors. The prevailing modeling strategies use the same current transport factor used for the definition of the α current gain in standard bipolar transistors [262, 274, 275]. As for the case of standard BJTs²⁵, HBTs are also formed by two anti-series P-N

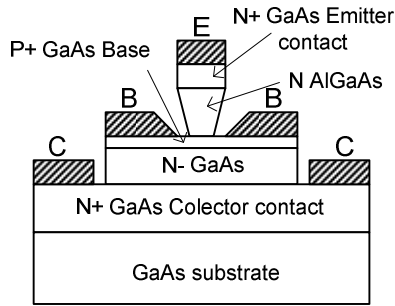


Figure 4.9 – GaAs HBT: construction detail.

junctions. The main differences are the type of semiconductors, and the doping profiles that are adequately adjusted within each region of interest (collector, base and emitter), as suggested by Figure 4.9. The delay nature arises due to the same high frequency considerations; and the base region must be modeled as a distributed transmission medium. The general model comprising delay influences on both base and collector regions can be

described using the α current gain (or alternatively, the transconductance parameter) as,

$$\alpha(j\omega) = \frac{\alpha_0}{1 + j\omega\tau_B} \frac{\sin(\omega\tau_C/2)}{\omega\tau_C/2} e^{-j\omega\tau_C/2} \quad (4.20)$$

where τ_B models the base transit time and τ_C the delay associated with the BC depletion layer.

The $\sin(x)/x$ term is often approximated by unity for frequencies one decade less than $2/\tau_C$, leaving a delay term plus one pole transfer function, as in (4.17).

4.3.2 Delays in Field Effect Transistors

Field effect transistors (FET) have some similar details with bipolar transistors. In both cases the transistor is formed by three semiconductor layers of alternating type. However, FETs have different current conduction mechanisms from their bipolar counterparts. Bipolar transistors rely on two P-N junctions to control currents, BC and BE that must be reverse and direct biased in the forward active mode. FETs use a different effect: the characteristics of a conductive channel between two of its terminals (source and drain), are controlled by the potential applied to a third terminal, gate, that acts isolated from it. Two types of FET devices can be considered: isolated gate devices, having the gate terminal electrically isolated from the channel (MOSFETs and HEMTs fall under this category); and junction FETs, where the gate terminal forms a semiconductor junction

²⁵ BJT – Bipolar Junction Transistor.

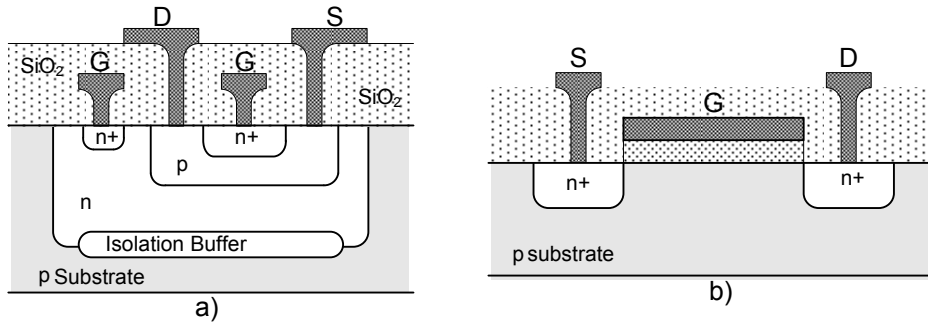


Figure 4.10 –Field effect transistor, construction details: a) JFET; b) MOSFET.

with the channel region (JFET and MESFET fall under this category). Figure 4.10 shows the construction details of the two basic FETs, the JFET and the MOSFET.

The current conduction mechanism in these two transistors is slightly different. Since current conduction occurs on the channel region, between the drain (D) and source (S) terminals, JFET devices are normally conductive while MOSFET devices are not. The reason for this distinction lies in the fact that JFET devices have a conductive channel between these terminals, while MOSFET devices have two opposite P-N junctions. The gate terminal acts in both cases as a current control input. For the JFET case, the gate junction must be reverse biased; the penetration of the depletion region under the gate junction, into the channel, constrains current conduction between the source and drain terminals. For the MOSFETs, the channel is usually made of type P semiconductor (or type N semiconductor, for P type devices). Since both drain and source are made of N+ semiconductor it is necessary to invert the channel type in order to have current flowing between these terminals. The process of channel inversion is explained by the capacitor like structure formed by the gate and the channel (where the gate and channel are the capacitor plates and the dielectric is formed by SiO_2). Applying a positive voltage to the gate terminal promotes the accumulation of positive charge on the gate plate. Since neutrality must be preserved, an equal amount of negative charge is attracted to the channel region under the gate (consisting of minority carriers on the P substrate), thus allowing the formation of a negative channel between drain and source.

MESFET and HEMT transistors have more complex construction details. As HBT transistors, these are generally fabricated with compound semiconductors. Also, as for the HBT case, MESFET and HEMT are high performance replicas of the simpler JFET and MOSFET transistors.

Figure 4.11 depicts the construction details of a MESFET, a conventional HEMT and a pseudomorphic HEMT (P-HEMT). The MESFET is similar to a junction FET except for the gate terminal; in MESFET transistors the gate terminal forms a metal-semiconductor junction with the channel, (also known as a Schottky junction [262, 274]). The principles of operation are similar to those described for JFET transistors.

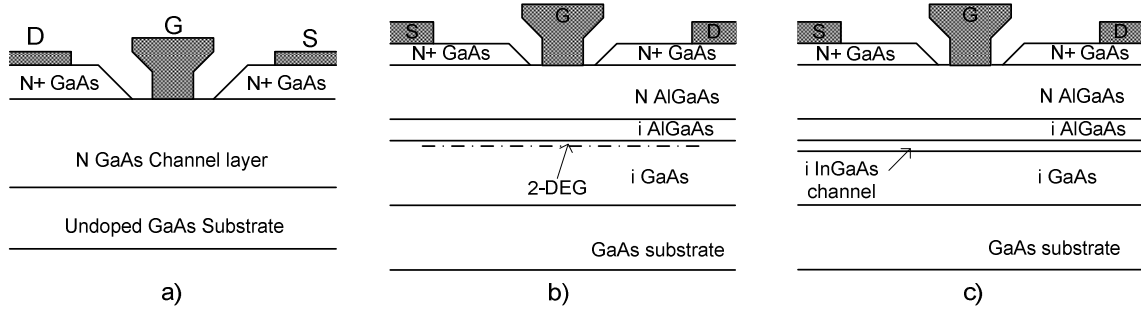


Figure 4.11 – High frequency GaAs FET transistors: a) MESFET, b) conventional HEMT, c) P-HEMT.

HEMTs involve more complex fabrication procedures. The various layers of semiconductor materials shown in Figure 4.11 are formed using modulation doped techniques. The essential characteristic of modulation doped semiconductors is the formation of a thin layer with enhanced conductive properties. In a conventional HEMT, the association of semiconductors with different band gaps originates a thin layer of accumulated free carriers (electrons). This layer is generally described as two-dimensional sheet of free carriers (2-DEG), which can move freely under the action of an externally applied electric field, without suffering from lattice scattering due to the recombination process [260, 262, 272, 274]. HEMT transistors use the 2-DEG layer as conductive channel. Current control is accomplished by the same mechanism as in JFET and MESFET transistors; the depletion region induced by the reverse biased gate junction constrains the channel conductive properties. However, since the 2-DEG layer is effectively separated from the gate junction, the HEMT structure resembles also a MOSFET transistor. P-HEMT devices use a high performance semiconductor compound on the place of the 2-DEG layer, leading to enhanced frequency performance devices.

Particularly important is that current conduction in FET devices is processed under the same type of semiconductor material, without involving any PN junction. However, both diffusion and drift current components still coexist. Drift currents arise due to the external voltage applied from drain to source. Assuming the source terminal as the reference, the PN junction near the drain terminal will be more reverse biased than the PN junction near the source terminal, implying that the depletion region is wider near the drain. This effect implies an asymmetric distribution of free carriers on the channel and consequently the existence of diffusion currents.

Nevertheless delay modeling for FET transistors is slightly simpler than in bipolar transistors. The main factors contributing to FET delays are: velocity saturation, recombination of carriers on the channel and the distribute nature of the channel. Usually the delay effects in FETs is modeled using two approaches: i) the distribute nature of the channel is modeled as an RC equivalent circuit branch between the gate and source terminals (represented by c_{gs} and r_i on the small signal equivalent of Figure 4.12); ii) the excess phase arising from model reduction is included on the

transconductance parameter (τ in Figure 4.12). The excess phase term includes three additive delay contributions [166]: the transit time of the carriers on the gate, the channel charging delay due to the distributed nature of the channel, and the drain delay due to the drift of free carriers under the high field depletion region under the drain.

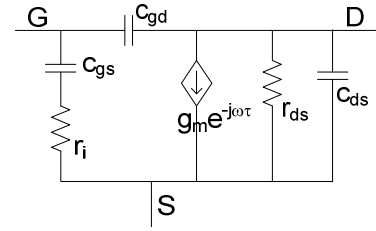


Figure 4.12 - FET small signal model comprising delay.

4.4 Delay Design

Chapter 3 discussed the effects of delays in feedback amplifiers. In general, these delays contribute to reduce the stability margins of closed-loop amplifiers. Nevertheless, it was shown that it is possible to explore these delays as a means of achieving significant bandwidth improvements, without seriously impairing the stability of the closed-loop amplifier. This section addresses the problem of controlling the net delay in delayed feedback amplifiers. Generally speaking, controlling delays resort to the design of suitable delay elements that can be incorporated within the amplifier. There are several possibilities to design these delay elements, namely: using properly terminated transmission lines [182, 183, 186, 188, 190, 191, 255, 261, 269]; using artificial transmission lines [201, 203, 205, 208, 211, 266] (often encountered in high frequency distributed delay equalizers); using all-pass filters with adequate phase modeling [195, 202, 209, 210] (Padé approximants of the complex exponential fall into these design category [198, 206, 207, 299]); or even using active devices or properly design active delay circuits [198-200, 204, 206, 207, 298-299]. Regardless of the technique, it must meet the stringent requirements of introducing small power, noise and area penalties on the overall amplifier's design. The following discussion will address the design of these delay elements, considering both active and passive circuits.

4.4.1 General Delay Representation

Delays are present in all electronic devices. The physical reason for this lies in the fact that electrical currents consist on charge moving from one point to another within the circuit. Since electrons cannot move with unlimited velocity, the charge transfer processes take a certain time. Transistors are no exception to this rule. As it was discussed in section 3.3, modeling delays in transistor is a complex task. Nevertheless it has become widely accepted that the effect of delays in transistors can be modeled within the transconductance parameter (or current gain). Since the transconductance of all the transistors within the amplifier will in general appear as a multiplying factor on the return ratio (especially when the direct gain is negligible), the effect of the intrinsic delays due to each transistor adds together. This observation suggests that a general modeling

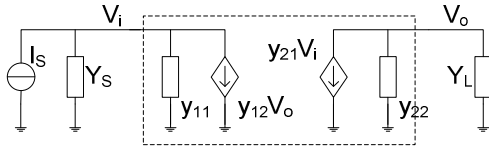


Figure 4.13 – General y parameter representation of the delay element.

strategy for delay elements within the amplifier should follow closely the transconductance example of a single transistor, using an equivalent y parameter circuit representation. Figure 4.13 shows the y parameter general representation of a delay element.

The load and source admittances were also included in order to account for loading effects. Using this representation it is possible to express any sort of transfer function between input and output. For instance, the voltage transfer between output V_o and input V_s (considering the Thévenin equivalent of the current source) is given by,

$$A_v = \frac{V_o}{V_s} = \frac{-y_{21}Y_s}{(y_{11} + Y_s)(y_{22} + Y_L)} \left[1 - \frac{y_{12}y_{21}}{(y_{11} + Y_s)(y_{22} + Y_L)} \right]^{-1} \quad (4.21)$$

Other gain relations can be expressed proportionally to A_v , namely: transadmittance $G_m = Y_L A_v$, transimpedance $Z_m = A_v / Y_s$ and current gain $A_i = Y_L A_v / Y_s$. For all transfer functions, the element y_{21} appears as a proportional factor. If the net delay is represented in y_{21} , all the transfer functions will experiment the same delay effect as A_v . The term inside brackets in (4.21) represents the internal feedback due to feedback paths within the delay element, represented by y_{12} . Usually this term can be detrimental to the overall circuit performance. However, for the majority of the applications it is possible to minimize internal feedback, using isolating buffers at the input and output of the delay element. It is interesting though to observe that the factor inside brackets on (4.21), provides the same bandwidth gain mechanism as a delayed feedback amplifier. Nevertheless, controlling the delayed feedback dynamics proper of a single transistor is not a realistic circuit designer task.

As discussed on the previous chapter, the delay can be represented by pure complex exponentials or any suitable frequency domain approximation of the complex exponential. For the first case, y_{21} will exhibit frequency dependence with transcendental character, that is,

$$y_{21}(s) = y_{21}^L(s) e^{-sL} \quad (4.22)$$

where $y_{21}^L(s)$ represents the frequency dependency of y_{21} without delay²⁶. For the second case, the simplest approach is to consider the 1st order Padé approximation of the complex exponential, and y_{21} will represent the RHP zero alone,

$$y_{21}(s) = y_{21}^T(s)(1 - sT) \quad (4.23)$$

where $y_{21}^T(s)$ represents the frequency dependency of y_{21} excluding the RHP zero. The complete Padé approximant can be synthesized using $y_{11} + Y_s$ or $y_{22} + Y_L$ as will be discussed in a later section.

²⁶ Obviously, this definition excludes the effect of group delay due to any phase term in $y_{21}^L(s)$.

4.4.2 Transmission Line Design

Transmission lines are the simplest and most immediate alternative to the synthesis of delay elements. Since electrons move with limited velocities, controlling the path length is equivalent to controlling the time delay needed to transport current between two circuit nodes. However, this simplistic approach is only valid under certain stringent constraints. The reason for these constraints is that every conductive medium between two circuit nodes is in fact a transmission line, having its associated transport mechanisms. These transport mechanisms arise due to the behavior of the electromagnetic fields on the transporting medium. A suitable classification considers the direction of the electric and magnetic fields against the propagating direction [255, 258, 269]. Under these classification transmission lines can be: TEM transmission lines (transverse Electric and Magnetic, both transverse to the propagation direction) – usual in every two conductor configurations; Higher mode transmission lines (where one of the electric or magnetic fields have a component on the propagation direction) – hollow single conductor wave guides and optical fiber fall into this category; and TEM space waves – when the transmission medium is the open space.

From a circuit design perspective, the most interesting is the first type, TEM two conductor transmission lines. For this kind of transmission lines it is possible to devise an R-L-C-G distributed model as depicted on Figure 4.14 [255, 258]. This model describes the transmission line as a collection of incremental length cells (Δx) having associated distributed resistance, inductance, conductance and capacitance. The distributed resistance (R) and inductance (L) of the line are represented by the series branch, while the parallel branches represent the dielectric losses (G) and capacitance (C). Assuming that the voltage and current at the beginning of the incremental cell at length x are represented by $V(s, x)$ and $I(s, x)$ respectively (where the variables s and x stand for complex frequency and space dependencies, respectively), and assuming also that the transmission line is infinite and uniform (without bends and with uniform dielectric properties), then the voltage and current increments at the ending of the incremental cell, $x + \Delta x$ can be written as,

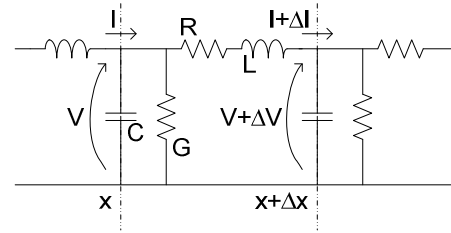


Figure 4.14 – Transmission line R-L-C-G distributed model.

$$\Delta V(s, x) = -I(s, x)(R + sL)\Delta x = -Z(s, x)I(s, x)\Delta x \quad (4.24)$$

$$\Delta I(s, x) = -V(s, x)(G + sC)\Delta x = -Y(s, x)V(s, x)\Delta x \quad (4.25)$$

where $Z(s, x)$ and $Y(s, x)$ represent the series branch impedance and the parallel branch admittance, respectively. Recalling that the uniformity assumption implies that both $Z(s, x)$ and

$Y(s, x)$ have no dependency on x , results on the well known Telegrapher's equations.

$$\frac{\partial^2}{\partial x^2} V(s, x) = -Z(s)Y(s)V(s, x) \quad (4.26)$$

$$\frac{\partial^2}{\partial x^2} I(s, x) = -Z(s)Y(s)I(s, x) \quad (4.27)$$

There are several different solutions for the Telegrapher's equations, depending on the particularities of the line. Considering the uniform case, represented by (4.26) and (4.27), a general solution for the voltage along the line is,

$$V(s, x) = V_1(s)e^{-\gamma(s)x} + V_2(s)e^{\gamma(s)x} \quad (4.28)$$

where $\gamma(s)$ is the propagation constant given by

$$\gamma(s) = \sqrt{Z(s)Y(s)} \quad (4.29)$$

The first term in (4.28) represents an incident wave (traveling in the positive x direction), while the second represents a reflected wave (traveling in the negative x direction). Thus, according to (4.28) the voltage along the line is represented as the sum of two waves propagating in opposite directions along the line. The current along the line is found from (4.27) and (4.29)

$$I(s, x) = \frac{V_1(s)e^{-\gamma(s)x} - V_2(s)e^{\gamma(s)x}}{Z_o(s)} \quad (4.30)$$

where $Z_o(s)$ represents the characteristic impedance, given by

$$Z_o(s) = \sqrt{\frac{Z(s)}{Y(s)}} \quad (4.31)$$

Once the characteristic impedance and the propagation constant are known, it is possible to describe the line behavior using equations (4.28) and (4.30). In general, the uniformity assumption fails to hold. On such cases, both Z_o and γ become complex entities depending on both frequency and space. It is generally difficult to find suitable expressions for Z_o and γ depending on the particular configuration of the transmission line in this case. Cohn and Wheeler, among others, developed a whole set approximate equations to find Z_o and γ [174-180, 184, 185, 261]. Presently, it is often desirable to recur to electromagnetic simulators such as the ones furnished within the ADS (Advanced Design System from Hewlett Packard) or MicroWave Office (from AWR – Applied Wave Research). Nevertheless, it is instructive to have a clear picture over the limiting factors arising in transmission lines.

4.4.2.1 The Propagation Constant

The propagation constant given by (4.29) can be further divided into real and imaginary parts. Restricting the complex frequency variable s to the imaginary axis, using $s=j\omega$,

$$\begin{aligned}\gamma(j\omega) &= \alpha(\omega) + j\omega\beta(\omega) \\ &= \omega\sqrt{LC} \sqrt{\left(1 + \frac{R^2}{\omega^2 L^2}\right) \left(1 + \frac{G^2}{\omega^2 C^2}\right)} e^{j\frac{\theta_\gamma}{2}}\end{aligned}\quad (4.32)$$

where θ_γ represents the phase of $Z(j\omega)Y(j\omega)$; $\alpha(\omega)$ is the attenuation constant, given by the real part of $\gamma(j\omega)$; and $\beta(\omega)$ is the phase constant, given by the imaginary part of $\gamma(j\omega)$. In general both $\alpha(\omega)$ and $\beta(\omega)$ are frequency dependent parameters. Furthermore, if the line is not uniform these parameters may also depend on space coordinates.

R and G characterize the line losses, where R represents the conductor resistivity and G the dielectric properties. For the ideal case, the lossless line, both R and G are zero, and the propagation constant becomes purely imaginary. That is, the attenuation is zero and the phase constant becomes $(LC)^{0.5}$. Inserting this simplification into (4.28), the voltage at any arbitrary point in the line becomes,

$$V(j\omega, x) = V_1(j\omega)e^{-j\omega x\sqrt{LC}} + V_2(j\omega)e^{j\omega x\sqrt{LC}} \quad (4.33)$$

Equation (4.33) is the Fourier transform corresponding to the sum of two voltage waves shifted in time, $v_1(t - x(LC)^{0.5}) + v_2(t + x(LC)^{0.5})$. This observation clearly explains the delay nature of a transmission line. Furthermore, from a delay design perspective the transmission line must be as close as possible to the lossless case. It is interesting to see that the factor $x(LC)^{0.5}$ in (4.33) represents the product of a distance (x) by the inverse of a velocity $((LC)^{0.5})$, thus the phase velocity can be defined accordingly as $v_p = (LC)^{-0.5}$.

For the general case, when both R and G have non-zero values, the expressions for the phase and attenuation result from (4.32). If both $\alpha(\omega)$ and $\beta(\omega)$ are made frequency independent [255], equation (4.33) still holds, except for two factors dependent on the attenuation. When such independence is difficult to attain, the delay terms in (4.33) become frequency dependent and the line presents dispersion²⁷. There are several factors affecting the line attenuation. The attenuation can be divided into three components: i) due to the conductor resistivity, increasing with the frequency due to the skin effect [174, 255, 261, 269]; ii) due to dielectric with finite resistivity, these are especially relevant for the case of integrated circuit transmission lines, where the dielectric

²⁷ Dispersion implies that signal components with different frequencies experiment different delays.

may be a semiconductor material; and iii) due to radiation losses, when power is radiated to the free space or to another neighbor line (mainly due to line ends and discontinuities).

4.4.2.2 The Characteristic Impedance

Following the above approach, the general expression for the characteristic impedance is given after (4.31) as,

$$Z_o(j\omega) = \sqrt{\frac{L}{C}} \sqrt{\left(1 + \frac{R^2}{\omega^2 L^2}\right) \left(1 + \frac{G^2}{\omega^2 C^2}\right)^{-1}} e^{j\frac{\theta_Z}{2}} \quad (4.34)$$

where θ_Z represents the phase of $Z(j\omega)/Y(j\omega)$. For the lossless case, $Z_o(j\omega)$ is given by $(L/C)^{0.5}$. When the line losses are relevant, the characteristic impedance becomes a complex function. The high frequency behavior of the transmission line resembles the ideal lossless case: as frequency increases, the loss contributions due to R and G become negligible both in $Z_o(j\omega)$ and $\gamma(j\omega)$.

4.4.2.3 General Transmission Line Delay Model

It is possible to find an adequate y parameter representation of the transmission line using equations (4.28) and (4.30). Consider the test set-up of Figure 4.15, where a transmission line with characteristic impedance Z_o and propagation constant $\gamma(s)$ and length l , is used to connect a load impedance Z_L to a generator V_s having internal impedance Z_S . In order to find the y parameter representation of this transmission line, two short-circuit test are made: i) shorting the load and taking $Z_S=0$, find the parameters y_{11} and y_{21}

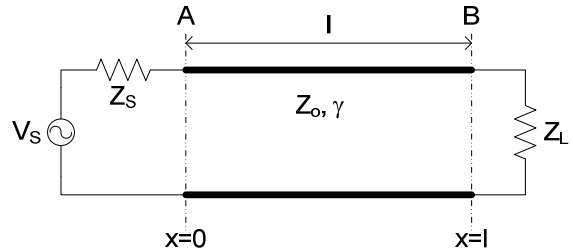


Figure 4.15 – Reference transmission line set-up.

representing the input admittance of the line and the transadmittance from port A (generator end, $x=0$) to port B (load end, $x=l$), respectively; and ii) shorting the generator port and applying a test generator instead of the load, find the parameters y_{22} and y_{12} representing the input admittance of the line and the transadmittance from port B to port A, respectively. Clearly, the above procedure leads to the equivalences $y_{11}=y_{22}$ and $y_{12}=y_{21}$, because of the reciprocity of the line. Pursuing with these results in the following set of y parameters,

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \frac{1}{Z_o} \begin{bmatrix} \coth(\gamma(s)l) & -\frac{1}{\sinh(\gamma(s)l)} \\ -\frac{1}{\sinh(\gamma(s)l)} & \coth(\gamma(s)l) \end{bmatrix} \quad (4.35)$$

According to the general delay model introduced in section 4.4.1, the delay effect can be best understood analyzing the voltage ratio between the voltage at port B against the voltage at port A. Considering the impedances Z_o , Z_L and Z_s and applying the voltage gain formula (4.21) to the circuit set-up of Figure 4.15, with the y parameters of (4.35), results in

$$A_v(s) = \frac{V_B(s)}{V_A(s)} = \frac{Z_o}{Z_o + Z_s} \frac{(1 + \rho_L) e^{-\gamma(s)l}}{1 - \rho_L \rho_s e^{-2\gamma(s)l}} \quad (4.36)$$

where ρ_L and ρ_s represent respectively the load and generator reflection coefficients, given by,

$$\begin{aligned} \rho_s &= (Z_s - Z_o) / (Z_s + Z_o) \\ \rho_L &= (Z_L - Z_o) / (Z_L + Z_o) \end{aligned} \quad (4.37)$$

Equation (4.36) represents the transmission line transfer function under arbitrary load and generator terminations. The utility of (4.36) can be best appreciated under certain generator/load constraints. Table 4.1 presents several generator load termination conditions and the correspondent voltage transfer function according to (4.36). There are three special termination conditions: i) exact matching, when the termination impedance equals the characteristic impedance; ii) short-circuit, when the termination is an short-circuit; and iii) open-circuit, when the termination is an open-circuit. Clearly the generator end of the line must be connected to the rest of the circuit, thus excluding the open-circuit condition. Similarly, the load end is always represented by some non-zero impedance, thus excluding the short-circuit condition. Analyzing all the load and generator conditions (except the two aforementioned cases and the general case represented by (4.36)), reveals that the voltage transfer is always proportional to the factor $e^{-\gamma(s)l}$, except for the case when the generator is a short-circuit and the load an open-circuit. These special generator/load configurations suggest possible circuit implemenations of a delay element able to produce a time delay of $\gamma(s)l$.

Table 4.1 – Delay transfer functions.

ρ_s	ρ_L	$A_v(s)$
0	0	$\frac{e^{-\gamma(s)l}}{2}$
0	1	$e^{-\gamma(s)l}$
0	ρ_L	$\frac{(1 + \rho_L) e^{-\gamma(s)l}}{2}$
-1	0	$e^{-\gamma(s)l}$
-1	1	$\frac{1}{\cosh(\gamma(s)l)}$
ρ_s	0	$\frac{Z_o}{Z_o + Z_s} e^{-\gamma(s)l}$

Considering the lossless case, equation (4.36) becomes,

$$A_v(j\omega) = \frac{Z_o}{Z_o + Z_s} \frac{(1 + \rho_L) e^{-j\omega\beta l}}{1 - \rho_L \rho_s e^{-2j\omega\beta l}} \quad (4.38)$$

Two important features of (4.38) are the magnitude function and the associated delay. Defining the delay as the symmetric of the first frequency derivative of the phase term correspondent to (4.38), results in

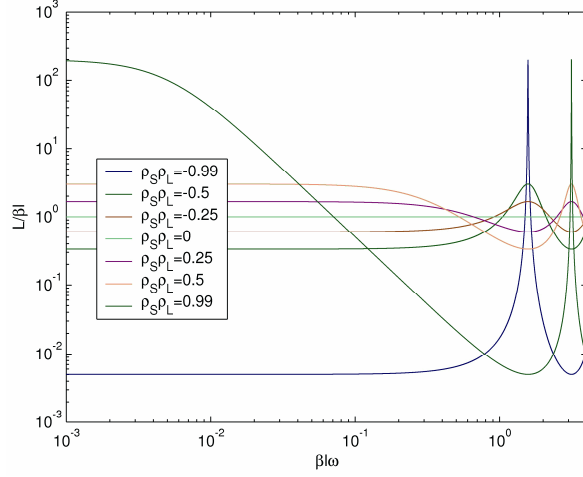


Figure 4.16 – Frequency dependence of the delay according to (4.39).

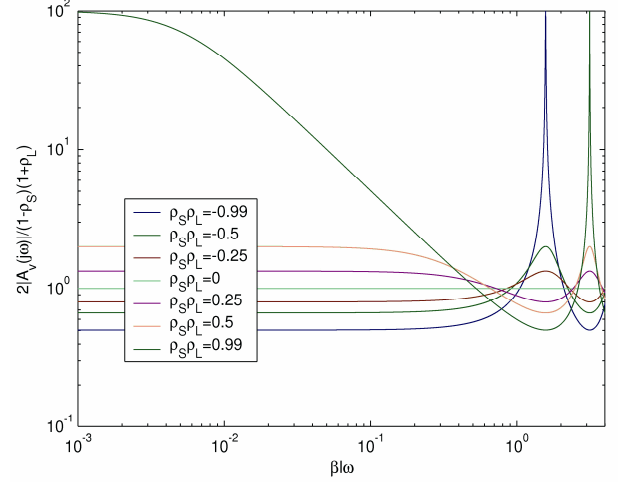


Figure 4.17 – Magnitude of the frequency response of the voltage transfer function (4.38).

$$\begin{aligned}
 L(\omega) &= -\frac{\partial}{\partial \omega} \arg[A_v(j\omega)] \\
 &= \frac{(1 - \rho_s^2 \rho_L^2) \beta l}{1 + \rho_s \rho_L (\rho_s \rho_L - 2 \cos(2\beta l \omega))}
 \end{aligned} \tag{4.39}$$

Figure 4.16 and Figure 4.17 represent the frequency dependence of the effective delay according to (4.39) and the magnitude of (4.38) for various values of $\rho_s \rho_L$. As can be seen, both delay and magnitude of the voltage transfer function become strongly frequency dependent for situations departing from the ideal matched case (one or both ends matched). The extreme case where both ends are opened corresponds to an extreme frequency dependency on both magnitude and delay functions. On the other hand, having the generator end shorted and the load end opened, represents a useful situation, since both delay and magnitude functions are almost flat for normalized frequencies less than unity. However, the effective delay for this situation is quite different from the ideal βl value.

s parameters are another elegant alternative for the representation of the transmission line as a set of two port parameters. On this work the y parameter representation is the preferred approach, since y parameters are closely related to the admittance matrix arising from circuit's nodal analysis. Nevertheless, from a measuring perspective, it is useful to consider the s parameters instead. It suffices to remember that y parameters are obtained using short circuit tests on the relevant ports of the circuit network, which are difficult to realize experimentally, while s parameters simply require referenced terminated tests to evaluate the circuit's frequency behavior. It is possible to convert y parameters into s parameters and vice-versa [251, 255, 261, 269]. For a transmission line under the set-up of Figure 4.15, maintaining the same assumptions made for the derivation of (4.35), the s

parameters are given by

$$\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \frac{1}{1 - \rho_s \rho_L e^{-2\gamma(s)l}} \begin{bmatrix} -\rho_s + \rho_L e^{-2\gamma(s)l} & \sqrt{(1 - \rho_s^2)(1 - \rho_L^2)} e^{-\gamma(s)l} \\ \sqrt{(1 - \rho_s^2)(1 - \rho_L^2)} e^{-\gamma(s)l} & -\rho_L + \rho_s e^{-2\gamma(s)l} \end{bmatrix} \quad (4.40)$$

Equation (4.40) reveals directly the importance of having the matched condition on both ports when synthesizing a required delay. In this case, equation (4.40) becomes

$$\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} 0 & e^{-\gamma(s)l} \\ e^{-\gamma(s)l} & 0 \end{bmatrix} \quad (4.41)$$

4.4.2.4 Microstrip and Stripline Transmission Lines

Figure 4.18 depicts the two most common planar two conductor transmission lines configurations: microstrip and stripline. Both configurations comprise a conductor used to convey

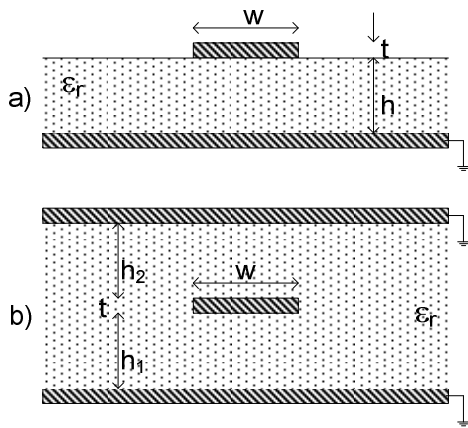


Figure 4.18 – Planar two conductor transmission lines: a) microstrip; b) stripline.

signals; a dielectric material providing the physical support of the structure; and one or two ground planes depending on each of the configurations, microstrip or stripline.

Stripline was the first reported planar topology, invented by Wheeler as a planar extension of coaxial cables [261]. In a stripline configuration, both electric and magnetic fields originated by the signals flowing on the central conductor are confined between the two ground planes, as in a coaxial cable. Cohn and Wheeler developed a set of approximated formulas for the

computation of both characteristic impedance and propagation constant [174-176, 185]. These formulas depend on several parameters, namely: the line geometry, dielectric properties (dielectric permittivity and conductivity), conductor losses (conductivity and skin effect), and magnetic permeability. A well organized collection of these results can be found in [261].

The extension of stripline to microstrip was quite obvious. Since these planar topologies were finding applications in printed circuit boards (PCB), the microstrip case represented an economy of conductive layers (only two necessary, while stripline require three conductive layers). Wheeler provided the first insight over the microstrip design problem [179, 180, 184]. One particular problem in microstrip is that the field lines cross two dielectrics with very different intrinsic properties: the supporting material and open-air. This slight difference turns the microstrip

modeling problem into a complicated one. Computational methods are the most common approaches to deal with microstrip lines [181, 187, 189, 261]. Two important assets in these two topologies is the fact that both characteristic impedance and phase velocity are normalized to the open-air case:

$$\begin{aligned} c &= \frac{1}{\sqrt{\mu_o \epsilon_o}} = 2.9979 \times 10^8 \text{ m/s} \\ \eta_o &= \sqrt{\frac{\mu_o}{\epsilon_o}} = 120\pi \text{ } \Omega \end{aligned} \tag{4.42}$$

where c is the light velocity and η_o the characteristic impedance on the open-air. These normalizations stem from the fact that both electrical permittivity and magnetic permeability are normally expressed as functions of ϵ_o and μ_o (permittivity and permeability of the open-air, 8.85pF/m and 1.26nH/m respectively).

Both stripline and microstrip configurations can be used in integrated circuits. Taking as reference the CMOS or BiCMOS SiGe 350nm processes from AMS (Austria Micro Systems), both comprising four metal layers and SiO_2 as dielectric, four implementation problems arise:

- i) The area occupied by the delay line is a serious compromising factor. Fortunately there are suitable techniques to reduce area while maintaining the necessary length, for instance, using spiral or meander paths.
- ii) Recurring to spiral or meander paths represent another complication factor. In both cases, the line is composed by multiple parallel strips of metal, which imply electromagnetic coupling phenomena. In general, coupling effects degrade the line performance, making the propagation constant and the characteristic impedance space dependent, implying large dispersion for higher frequencies.
- iii) Since the height and electrical properties of dielectric and metal layers are imposed by the technology, the line width is the only disposable design parameter to establish the characteristic impedance. The characteristic impedance is always less than η_o , and displays an inverse dependence on the line width. These observations are especially critical from the integrated circuit design perspective: first because high characteristic impedances require smaller line widths and second because small characteristic impedances imply large currents. Since the maximum current density of the metal layers fix the minimum possible line width, these two conditions are difficult to manage. The preferable design strategy is to use microstrip topologies designed on the top metal layer, which generally have thick metal with enhanced conductivity and maximum current density.

- iv) Finally, since these problems involve high complexity modeling procedures, it is usually preferred to resort to electromagnetic simulators, usually not available within the dominant integrated circuit software design packages (like Cadence Design Frame Works, Mentor Graphics and Tanner Tools).

4.4.2.5 Meander Delay Lines

Figure 4.19 shows the typical structure of a meander delay line. Meander lines are occasionally used to design integrated resistors (using suitably chosen resistive layers), inductors (controlling

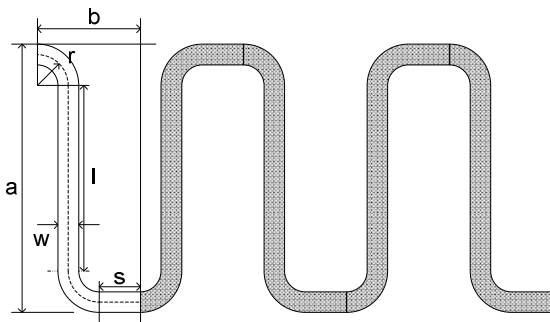


Figure 4.19 – Meander line.

appropriately the aperture between stripes and the stripe width and length) and delay lines (controlling the coupling between stripes). Meander lines allow important area/length benefits, since for a fixed length the occupied area can be adequately set once the stripe separation and length are chosen.

From a delay perspective, meander lines should reflect as close as possible the behavior of a simple straight and matched transmission line. This is however perturbed due to two main factors: the line bends and the stripe separation. Meander lines are two dimensional structures having electric and magnetic fields that interfere differently for different sections of the line. This interference phenomenon is known as coupling. In a broader sense, coupling between two parallel lines arise due to electrical currents induced on one line due to the electromagnetic field distribution on the other. Coupling problems are difficult to analyze without an electromagnetic simulator. Nevertheless several results have been reported focusing on both theoretical and experimental evidence of coupling effects in meander structures. In particular: i) Cohn and Wheeler addressed the problems of coupling in microstrip and stripline configurations, producing a set of approximate formulas [176, 178, 184], which can be found in an augmented and organized form in [261]; ii) applications of meander lines as delay equalizers can be found in [182, 183] and the references therein; iii) modeling strategies suitable for delay element synthesis in meander lines were also reported [186, 188, 190, 191].

Accordingly to these results, there are two effective methods which can be employed to reduce coupling: i) using shielded lines, thus confining the electromagnetic fields; and ii) increasing the separation between the interfering lines, the preferred method due to its simplicity. It is commonly accepted that for stripe separation larger than four times the stripe width the coupling effects become increasingly negligible and the meander line structure approaches the behavior of a simple

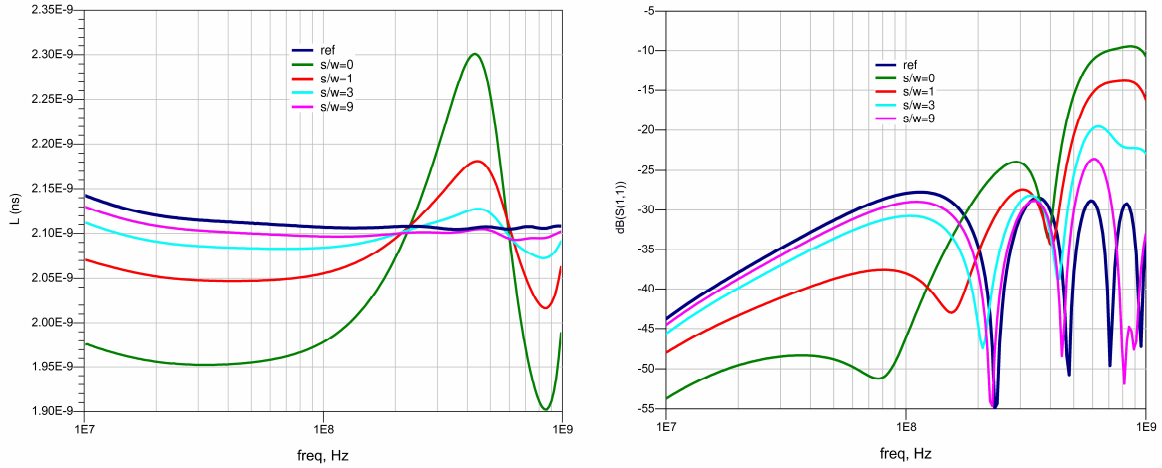


Figure 4.20 – Meander line delay design example: effective delay (L) and input reflection coefficient (S_{11}).

straight line having the same middle section length [188, 190, 191] (except for the bend contributions).

In order to have a clear understanding of the coupling phenomena in meander lines, some electromagnetic simulations were carried out with ADS, on a reference delay line having a topology similar to the one described in Figure 4.19. Figure 4.20 depicts electromagnetic simulation results for a meander delay line with effective length of 337mm and 1.4mm width on a cooper FR4 substrate (having: $\epsilon_r=4.7$, $\mu_r=1$, $h=0.8\text{mm}$, $t=35\mu\text{m}$, $\sigma=5.76\text{MS/m}$, $\tan\delta=0.014$). This simulation example used a meander line configuration like the one depicted in Figure 4.19, having three stripes, with $r=w$ and s/w as design parameter (accounting for the stripe separation²⁸). The stripes used a fixed length of $l=100\text{mm}$. It is clear that the delay varies with frequency for all the tested situations, showing that the line is dispersive. Also shown is the effect of stripe separation compared against the reference case (a straight line with the same length, represented by the blue line), showing that wider separations effectively reduce coupling effects. The degradation of the delay due to coupling is patent on: i) the delay reduction – smaller delay correspond to lines having smaller stripe separation and thus suffering more from coupling interference; ii) delay variation – implying larger dispersion, also observed for smaller stripe separations; and iii) loss of matching – the input reflection coefficient (S_{11}), especially exhibits large degradation for smaller stripe separations. Meander lines with more stripes exhibit larger coupling effects due to the presence of more interfering electromagnetic field lines. Nevertheless, using stripe separations larger than 4 proves effective, confining coupling effects originating on one stripe to the two immediately adjacent stripes (and vice-versa).

For practical design purposes it is advisable to consider simple approximations for both

²⁸ Note that the stripe separation is for this case (with $r=w$) given by $s+w$.

characteristic impedance and line delay. Such simple approximations can be derived assuming a microstrip with zero thickness conductors (t in Figure 4.18) above the ground plane. It has been shown that this approximation holds well when the conductor thickness is negligible when compared to the dielectric height ($t \ll h$) [258, 261]. Table 4.2 presents a collection of such approximations [258, 261]. The first formula on Table 4.2 allows for relative permittivity correction due

Table 4.2 – Microstrip design formulas.

Design Formulas	
$\epsilon_{r_{eff}}$	$\frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left(1 + 10 \frac{h}{w}\right)^{-\frac{1}{2}}$
Z_o	$\frac{\eta_o}{2\pi\sqrt{\epsilon_{r_{eff}}}} \ln\left(8 \frac{h}{w} + \frac{w}{4h}\right) \Leftarrow \frac{w}{h} \leq 1$ $\frac{\eta_o / \sqrt{\epsilon_{r_{eff}}}}{w/h + 2.42 + 0.44 h/w + (1 - h/w)^6} \Leftarrow \frac{w}{h} \geq 1$
l_{eff}	$n(l + s + \pi r)$
A_{eff}	$n(l + w + 2r)(s + 2r)$

to the presence of two dielectrics in the microstrip structure. The next approximated formulas allows the computation of the characteristic impedance, holding respectively for small line widths ($w/h < 1$) and large line widths ($w/h > 1$). The final formulas are entirely based on the meander topology of Figure 4.19: the first holds for the middle section effective length (assuming n sections of equal lengths), and the second for the effective area (assuming n sections of equal area). The effective delay can be computed using the phase velocity and the effective length,

$$L = \frac{l_{eff}}{c} \sqrt{\epsilon_{r_{eff}}} \quad (4.43)$$

For the above example, the characteristic impedance and the effective delay are approximately, 55.4Ω and $2.107ns$ respectively, compared with the simulation results for the straight line at $1GHz$ these values have an error of 10% and less than 1% respectively.

Figure 4.21 depicts the area filling factor as a function of two design parameters, the normalized stripe length (l/w) and the normalized separation width (s/w). The area filling factor represents the area penalty taken against the effective area of the line. For a fixed separation width of 4, the filling factor varies between 3 and 4.5 when, l/w ranges from 1 to 10, meaning that the area penalty in choosing l/w is not very significant. However, if s/w increases, the area penalty becomes

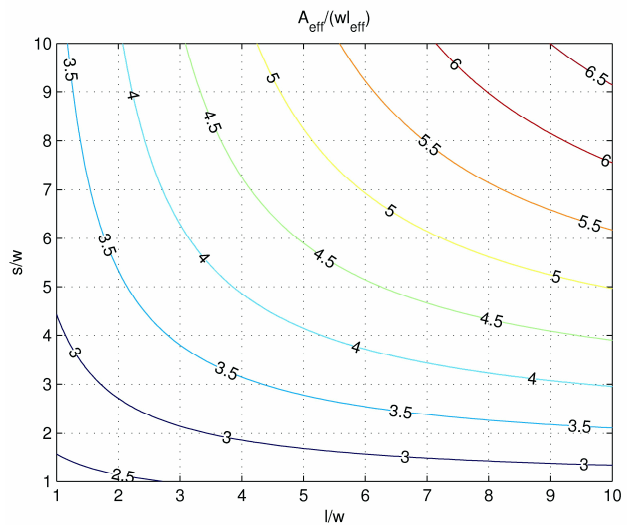


Figure 4.21 – Area filling factor.

strongly dependent on l/w . One final aspect regarding the area occupancy of the meander line structure deals with the microstrip line topology. The same assumptions made for the adequate stripe separation regarding coupling effects, are also valid regarding the fringing fields of the line boundaries. This means that the ground plane below the meander line should extend at least four times the width stripe from the line boundaries, thus preventing radiation losses due to fringing fields on these regions.

4.4.3 Active Delay Design

As discussed in section 4.1 delays in electronic systems are classified into two categories: intrinsic and designed. Section 4.3 presented a simplified discussion over the nature of intrinsic delays in transistors. Given the modeling complexity of these intrinsic delays, their effect is best taken into account during circuit optimization. There are several cases where such optimization procedures require to intentionally add extra delay elements to optimize circuit performance; such as the design of equalization filters required to combat inter symbol interference (ISI) [201, 205-208, 211, 212], clock synchronization using delay locked lines (DLL) [199, 200, 204], and the delayed feedback amplifiers here addressed. Some of these examples share the necessity of controlling the delay. For instance, bandwidth optimization procedures in delayed feedback amplifiers require delay control to cope with gain adjustments, as previously discussed in chapter 3. For such cases, using passive structures or even transmission line to synthesize delay elements represents a limiting factor, since the effective delay is then fixed. Furthermore, depending on the delay value, passive implementations may pose severe area penalties: i) large delays imply large transmission line lengths and consequently large area; ii) recurring to artificial transmission lines using distributed or lumped structures to approximate the all-pass characteristics of the ideal delay also requires a large implementation area [195, 201-203, 205, 206, 208, 209-212]. Active delays, on the other hand, furnish an attractive alternative design solution, combining both controllability and reduced area penalty [198-200, 204, 207, 298-299]. Active delays are preferably used in design applications where area represents an important design constraint, such as in integrated circuit design.

Active delays can be classified as digital or analog delays, according to the nature of the signals to process. Digital delay elements are generally encountered in digital circuits. Their operation is based on the saturated nature of digital signals. From binary logic point of view, a digital signal is a sequence of well defined signal voltages representing high (1) and low (0) logic levels. In its very essence, the digital signal is truly an analog signal, since the transitions between logic levels are not instantaneous. Digital delay elements explore these transitions as a means of retarding future

decisions. It is possible to implement these digital delay elements using a multitude of different methods. The simplest solution consists of a current starved inverter: since the transition time in an inverter is roughly proportional to the load capacitance and the average charge/discharge current, controlling the current through the inverter is a simple and effective method of controlling the delay. CMOS implementations of this concept are widely found in the literature [199, 200, 204].

Active analog delay elements are not as simple as their digital counterparts. There are two main reasons for this increased difficulty: i) first, because signals are treated in the analog domain and must as such, preserve their original form: the output signal should be as close as possible a delayed replica of the input; ii) second, because circuits consisting of transistors and/or passive R-L-C elements can only approximate the desired all-pass transfer function of the ideal delay. There are several methods to synthesize all-pass transfer functions able to approximate the ideal analog delay. These techniques are based on rational transfer functions comprising adequate pole-zero patterns [195, 202, 209, 209]. One important family of such approximations is the well known Padé approximants [268]. Padé approximants exhibit the desired properties of constant magnitude and small complexity. The following discussion will show that first order Padé approximants are also simple to synthesize using single transistor circuits.

4.4.3.1 First Order Padé Approximant Delay Modeling

Analog active delays can be modeled using first order Padé approximants of the form [268],

$$e^{-sT} \cong \frac{1-sT/2}{1+sT/2} = \Delta(s) \quad (4.44)$$

where T represents the time delay in seconds. In (4.44) it is immediately recognized the presence of a pole-zero pair, where the zero is placed on the RHP, with angular frequency symmetric to pole. The effective delay is given by the first frequency derivative of the argument of (4.44). Restricting the complex frequency s to the imaginary axis, this results in

$$\begin{aligned} T(\omega) &= -\frac{\partial}{\partial \omega} \arg \left[\frac{1-j\omega T/2}{1+j\omega T/2} \right] \\ &= \frac{T}{1+(\omega T/2)^2} \end{aligned} \quad (4.45)$$

Equation (4.45) shows that the effective delay is essentially constant for frequencies where $\omega \ll 2/T$. The implementation of (4.44) in transistorized circuits implies some uncertainties on the pole and zero magnitudes. These uncertainties are best viewed as pole-zero mismatches that lead to degradation of the delay effect. This degradation can be modeled using the following equation;

$$\Delta(s) = \left(\frac{2-s(T+\Delta_z)}{2+s(T+\Delta_z)} \right) \left(E_o \frac{2+s(T+\Delta_z)}{2+s(T+\Delta_p)} \right) = \Delta_{eff}(s) E(s) \quad (4.46)$$

where, Δ_p and Δ_z are respectively the pole and zero displacement frequencies, and $E_o=E(0)$ represents a constant gain due to mismatch. The term $\Delta_{eff}(s)$ is the effective delay and $E(s)$ accounts for the mismatch error (that affects both amplitude and phase). A better mismatch characterization consists of analyzing the maximum of the amplitude error, E_A , and the maximum of the phase error E_ϕ . Once again, taking $s=j\omega$, these error measures are defined as

$$E_A = \left| \frac{E_o - E(\infty)}{E_o} \right| = \left| \frac{(\Delta_p - \Delta_z)/T}{1 + \Delta_p/T} \right| \quad (4.47)$$

$$E_\phi = \left| \max \left\{ \tan^{-1} \left(\omega \frac{\tau + \Delta_z}{2} \right) - \tan^{-1} \left(\omega \frac{\tau + \Delta_p}{2} \right) \right\} \right| \quad (4.48)$$

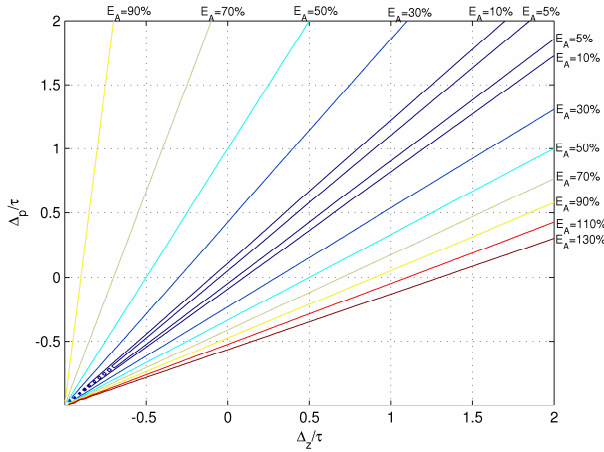
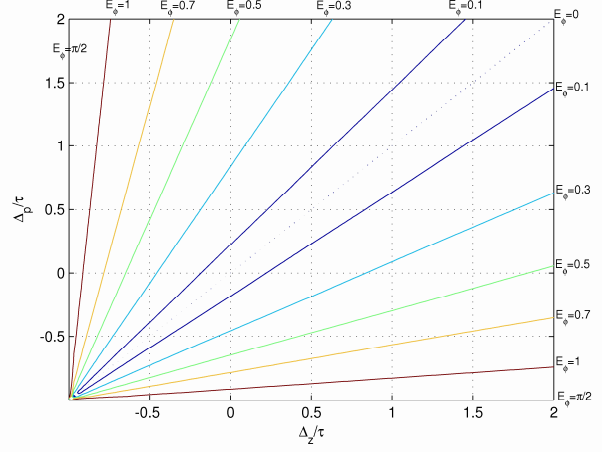
It is possible to eliminate the frequency variable from (4.48) by a simple evaluation of the zeros of the first derivative of E_ϕ , followed by a back substitution. The final result is,

$$E_\phi = \left| \tan^{-1} \left(\sqrt{\frac{1+\Delta_z/T}{1+\Delta_p/T}} \right) - \tan^{-1} \left(\sqrt{\frac{1+\Delta_p/T}{1+\Delta_z/T}} \right) \right| \quad (4.49)$$

E_A characterizes the height of the step like amplitude characteristic due to pole-zero terms in $E(s)$. On the other hand, E_ϕ represents the shifting of the phase characteristic. The associated delay error is frequency dependent. Following the same procedure to reach the result of equation (4.45), the delay error is given by,

$$\begin{aligned} \Delta_T(\omega) &= \frac{1}{2} \left(\frac{T + \Delta_z}{1 + \omega^2 (T + \Delta_z)^2 / 4} - \frac{T + \Delta_p}{1 + \omega^2 (T + \Delta_p)^2 / 4} \right) \\ &\approx \frac{1}{2} \frac{\Delta_z - \Delta_p}{1 + \omega^2 T^2 / 4} \end{aligned} \quad (4.50)$$

Equation (4.50) shows that the delay error is approximately proportional to the difference between the zero and pole displacement frequencies. Figure 4.22 and Figure 4.23 shows contour plots of the mismatch effects on $\Delta(s)$, using the amplitude and phase measures on (4.47) and (4.49). It can be seen that when $\Delta_p/T = -1$ or $\Delta_z/T = -1$, both E_A and E_ϕ attain their maximum values. However the behavior of E_A and E_ϕ is slightly different for these two cases. The phase error is symmetric on the both sides of $\Delta_p/T = \Delta_z/T$, the maximum phase shift occurs for $\Delta_p/T = -1$ or $\Delta_z/T = -1$ and has the same value, $\pi/2$ radians (or 90°). The same symmetry is not verified for the amplitude error. When $\Delta_z/T = -1$ the E_A reaches 100% of error – this is understandable, since for this value of Δ_z both the


 Figure 4.22 – Contour plots of the amplitude error E_A .

 Figure 4.23 – Contour plots of the phase error, E_ϕ .

RHP zero and the pole on $\Delta_{eff}(s)$ are effectively cancelled. On the other end, when $\Delta_p/T = -1$ the amplitude error increases without bound since the pole in $E(s)$ is nullified. For small values of E_A (less than 10%), the contour plot of Figure 4.22 is almost symmetric in the vicinity of $\Delta_p/T = \Delta_z/T$.

It should be observed that this modeling approach considered that the undesirable error term $E(s)$ has only first order dependencies on the frequency variable. Real circuit implementations of this concept may exhibit higher order terms. For such cases the above considerations are only first order approximations. Nevertheless the error measures E_A and E_ϕ can be defined in similar forms and measured from experimental data.

4.4.3.2 Transistorized Analog Delay based on First Order Padé Approximant

Analog implementations of the Padé approximant of the ideal delay can be found in several contributions [198, 207, 298-299]. In [198], Bult and Wallinga proposed a current-mode implementation of this Padé approximant based on a seven transistor current inverter. A simpler two transistor cell implementation was originally reported by Buckwalter and Hajimiri in [207]. This circuit is based on the transfer function of a degenerated common emitter stage, suitably

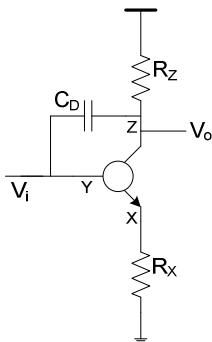


Figure 4.24 – Degenerated common X stage.

modified to approximate the desired Padé approximant. In [298-299] the Nero Alves *et al* extended and fully characterize all the parasitic contributions arising in these simple degenerated common emitter stages used as delay cells.

A simple implementation of the Padé delay consisting of only one transistor is depicted in Figure 4.24. It is commonly known the presence of a Miller capacitance in amplifying structures based on common X (CX) configurations (common

emitter (CE) and common source (CS) using Bipolar and MOSFET transistors respectively). The Miller effect is characterized by an effective multiplication of the input capacitance of the circuit, thus allowing for well known compensation strategies [267, 277, 280]. However a complete analysis of the circuit reveals also a zero on the RHP. Using an X degeneration resistor results in two effects: i) gain reduction, also implying large bandwidth (less input capacitance); and ii) equalization of the time constants (in absolute value) of the RHP zero and the left half plane (LHP) pole. These two effects allow a simple implementation of a Padé approximated delay. A complete analysis of the circuit using a generic transistor model comprising only input resistance, r_i , and a transconductance g_m (see appendix A 1), holds the following set of y parameters:

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} \frac{1+sR_Y C_D}{R_Y} & \frac{\beta-sR_Y C_D}{R_Y} \\ -sC_D & \frac{1+sR_Z C_D}{R_Z} \end{bmatrix} \quad (4.51)$$

where R_Y represents the input resistance seen from the terminal Y, given as $R_Y=r_i(1+\beta)R_X$, and β is the current gain ($\beta=g_m r_i$). Following the general delay modeling formulation of section 4.4.1, the delay effect is best represented by the voltage gain given by,

$$A_v = -\frac{y_{21}}{y_{22}} = -\beta \frac{R_Z}{R_Y} \frac{1-sR_Y C_D/\beta}{1+sR_Z C_D} \quad (4.52)$$

The time constants of the pole-zero pair are equal if the factor $\beta R_Z/R_Y$ is equal to unity. For this condition (4.52) resembles a first order Padé approximant, for which the effective delay would be $2R_Z C_D$ for frequencies $\omega < R_Z C_D$.

In general the circuit of Figure 4.24 has to be driven by a non-ideal source, thus the effect of non-zero output resistance of this source has to be accounted. This can be simply done considering the general delay model of Figure 4.13. Using the y parameters previously calculated and assuming that the load is simply R_Z , the modified y parameters referenced to the input source V_S (I_S/Y_S according to Figure 4.13) are given by

$$\begin{bmatrix} y_{11}^{R_S} & y_{12}^{R_S} \\ y_{21}^{R_S} & y_{22}^{R_S} \end{bmatrix} = \begin{bmatrix} Y_S + y_{11} & y_{12} \\ \frac{y_{21} Y_S}{Y_S + y_{11}} & \frac{\Delta_Y + y_{22} Y_S}{Y_S + y_{11}} \end{bmatrix} \quad (4.53)$$

where Δ_Y represents the determinant of the original y parameter matrix ($\Delta_Y=y_{11}y_{22}-y_{12}y_{21}$), and the superscript R_S ($R_S=1/Y_S$) denotes the R_S dependence. Using the modified y parameters, the voltage gain referenced to V_S becomes,

$$A_v^{R_S} = -\frac{y_{21}^{R_S}}{y_{22}^{R_S}} = -\beta \frac{R_Z}{R_{YS}} \frac{1 - s R_Y C_D / \beta}{1 + s \left[1 + \frac{R_S}{R_{YS}} \left(\beta + \frac{R_Y}{R_Z} \right) \right] R_Z C_D} \quad (4.54)$$

where R_{YS} is the input resistance including R_S ($R_{YS}=R_Y+R_S$). Applying the same unity gain condition as before, making $\beta R_Z=R_{YS}$, the zero and the pole are ruled by the time constant $T=R_Z C_D$ plus two displacement frequencies Δ_z^{RS} and Δ_p^{RS} defined as,

$$\begin{aligned} \frac{\Delta_z^{R_S}}{T} &= -\frac{R_S}{\beta R_Z} \\ \frac{\Delta_p^{R_S}}{T} &= \frac{R_S}{R_Z} \left(2 - \frac{R_S}{\beta R_Z} \right) \end{aligned} \quad (4.55)$$

As seen, the effect of R_S on the voltage transfer function of the circuit of Figure 4.24 traduces in a degradation of the pole-zero time constants and hence the degradation of the delay effect.

There are others sources of degradation in this circuit, namely: i) the intrinsic capacitances of the device, in particular c_x (between nodes Y and X), since c_z (between nodes Y and Z) together with C_D can define the effective delay; ii) due to layout considerations, there will always be some parasitic capacitances from nodes X, Y and Z to the ground; iii) insufficient detail of the assumed generic model, for instance the transistor's intrinsic delay may also contribute to the effective delay. The effects caused by these elements on the y parameters of the circuit of Figure 4.24 are easily evaluated, performing a series of adequate substitutions on the elements of (4.51). Considering the small signal equivalent circuit correspondent to the circuit of Figure 4.25, these substitutions are: i) $Z_X(s)=R_X/(1+sR_XC_{PX})$ instead R_X ; ii) $Z_Z(s)=R_Z/(1+sR_ZC_{PZ})$ instead R_Z ; iii) $z_i(s)=r_i/(1+sr_i c_x)$ instead r_i ; and $y_{II}(s)+sC_{PY}$ instead of $y_{II}(s)$. The effect of both intrinsic and parasitic capacitances on the effective delay can be evaluated using approximated analysis of the circuit in Figure 4.25. Table 4.3 summarizes these effects, using two approximations: i) first, the individual effect of each parasitic capacitance is computed; ii) second, the resultant transfer function is simplified considering only

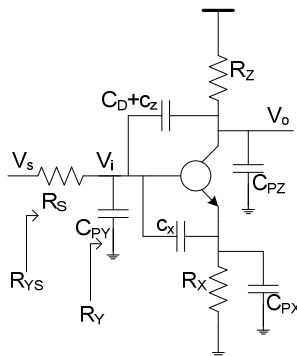


Figure 4.25 - Degenerated common X stage, with degenerating elements.

first order time constants. The third column of Table 4.3 presents the frequency restrictions due to the truncation of higher order terms. It is obvious that the number of terms of order superior to 1 increase when the parasitic capacitances are all accounted. However, for the purpose of circuit design the frequency of operation is much lower than the frequency associated to the second order time constants, which is approximately given by ω . Capacitance c_x is taken

Table 4.3 - Pole-Zero displacement due to parasitic elements.

Case	Δ_z	Δ_p	Frequency Restrictions $\omega_r^{-1} \ll \sum \frac{1}{\min\{\omega_i\}}$
R_S	$\frac{-R_S}{\beta}(C_D + c_z)$	$\left(2 - \frac{R_S}{\beta R_Z}\right) R_S (C_D + c_z)$	-
$R_S + c_x$	$\Delta_z^{R_S}$	$\Delta_p^{R_S} + \frac{R_S + R_X}{g_m R_Z} c_x$	$\omega_1 = \sqrt{g_m / R_X c_x (C_D + c_z)}$ $\omega_2 = \omega_1 \left[1 + R_S \left(\frac{1}{R_X} + \frac{1}{R_Z} \right) \right]^{-\frac{1}{2}}$
$R_S + C_{PX}$	$\Delta_z^{R_S} - R_X C_{PX}$	$\Delta_p^{R_S} + \frac{R_X (1 + R_S / r_i)}{g_m R_Z} C_{PX}$	$\omega_1 = \sqrt{g_m / R_X C_{PX} (C_D + c_z)}$ $\omega_2 = \omega_1 \left[1 + R_S \left(g_m + \frac{1}{r_i} + \frac{1}{R_Z} \right) \right]^{-\frac{1}{2}}$
$R_S + C_{PY}$	$\Delta_z^{R_S}$	$\Delta_p^{R_S} + \left(1 - \frac{R_S}{\beta R_Z} \right) R_S C_{PY}$	$\omega_1 = \left[(1 - R_S / \beta R_Z) R_S R_Z C_{PY} (C_D + c_z) \right]^{-\frac{1}{2}}$
$R_S + C_{PZ}$	$\Delta_z^{R_S}$	$\Delta_p^{R_S} + R_Z C_{PZ}$	$\omega_1 = \left[(1 - R_S / \beta R_Z) R_S R_Z C_{PZ} (C_D + c_z) \right]^{-\frac{1}{2}}$

here as a parasitic element, although its intrinsic nature (c_π for bipolar transistors (BJT) and c_{gs} for MOSFET).

Several observations can be made after a careful inspection of the results displayed on Table 4.3:

i) Both R_S and C_{PX} can cause same signal pole-zero frequency shifts, since both have similar effects on the pole-zero time constants. ii) All the other parasitic capacitances affect only the pole frequency, thus implying an immediate effect on bandwidth restriction. Their individual effects are nevertheless, inversely proportional to $g_m R_Z$ (except for C_{PZ}), thus allowing a general strategy for optimization. iii) Finally, the total contribution on the pole-zero displacement time constants can be evaluated by a simple addition of all the previous effects, as stated in (4.56),

$$\Delta_z = \Delta_z^{C_{PX}} \quad (4.56)$$

$$\Delta_p = \Delta_p^{C_X} + \Delta_p^{C_{PX}} + \Delta_p^{C_{PY}} + \Delta_p^{C_{PZ}} - 3\Delta_p^{R_S}$$

where the superscripts identify the individual contributions.

4.4.3.3 Degradation Effects on FET and Bipolar Circuits

As mentioned previously, the expressions in Table 4.3 were obtained using a generic transistor model. These effects have different manifestations for Bipolar and FET circuits. For Bipolar and HBT technologies the results in Table 4.3 are accurate as long as the transistor model resembles the generic model used for derivation purposes. For the FET case, the value of r_i is theoretically infinite, and this significantly reduces the contribution of the parasitic capacitances on the pole-zero displacements. Assuming that the unit gain condition is preserved, the displacement of the pole-zero

time constants due to R_S become 0 for the zero time constant, $2R_S(C_D + c_z)$ for the pole. Obviously, the pole displacement can be made negligible if R_S will be very small. Making the same analysis for the remaining parasitic capacitances, it can be observed that the zero displacement is always 0, except for the case of C_{PX} . The pole displacement becomes attenuated for the cases of C_X and C_{PX} , where the effective resistance appears divided by the factor $g_m R_Z$ ($R_S + R_X$ and R_X , respectively). The C_{PY} case, the pole displacement becomes controlled solely by R_S , thus allowing for simple minimization procedures.

The major sources of pole-zero time constants displacements are common to both Bipolar and MOSFET circuits. For the zero displacement C_{PX} is the dominant parasitic effect, while for the pole case C_{PZ} adds a major contribution.

The previous analysis seems to suggest that MOSFET are more efficiently used as delay elements than Bipolar Transistors. However this is in general false. The main reason is that MOSFET transistors have smaller g_m than their Bipolar counterparts. In order to achieve unit gain, the factor $g_m r_i R_Z / R_{YS}$ needs to be approximately equal to unity. Using MOSFET transistors this is accomplished increasing both R_Z and R_X , thus implying a magnification of the time constants. This effect reduces drastically the bandwidth of the delay cell. A meaningful estimate of the bandwidth can be computed using the results reported on [197]. Following this result, the bandwidth of the delay cell is given approximately as,

$$B \approx 2 / (\Delta_p - \Delta_z) \quad (4.57)$$

Note that the bandwidth estimate on (4.57) makes sense even for the case of $\Delta_p = \Delta_z$. When the pole and zero displacements are similar, the predicted bandwidth is very large. This is consistent with the fact that, for this situation, the term $E(s)$ may not even exhibit a 3dB cut-off frequency. Generally the displacement for the pole time constant is larger than for the zero, resulting always in positive values of bandwidth.

4.4.4 Active Delay versus Transmission Lines in Integrated Circuits

The last sections presented two strategies suitable to the design of delay elements. Several aspects regarding integrated circuit design of these delay elements deserve attention. First of all, due to the nature of the exposed procedures and to the presence of several degenerating factors, these delay elements can only approach the ideal delay behavior, that is, constant phase decay and unitary transfer gain. Nevertheless, it is possible to closely approach this behavior using the theoretical guidelines deployed in sections 4.4.2 and 4.4.3.

Two factors common to both strategies (active and passive design approaches) are the

degenerating effects caused by load and source considerations. For the transmission line, it is imperative to have adequate matching conditions at least at the load end of the line. It is also advisable to have low source impedance (lower than the characteristic impedance) in order to minimize loss due to voltage division between source and line. Considering now the case of the degenerated CX stage Padé approximated delay, the presence of non-zero source impedance on circuit's performance originates two effects: phase errors (implying delay errors) and amplitude errors. On the output port, a load impedance connected in parallel with R_Z , comprising both resistance and capacitance, imply a drastic reduction of the delay time constant ($R_Z(C_D + c_z)$).

From these perspectives, both types of delay elements should be adequately isolated from all neighboring circuitry. One possibility is to introduce two isolating buffers: one to reduce the source impedance and other to reduce the effects of loading presented to the delay element. The usage of these buffers is especially mandatory for the transmission line case, where the source impedance is preferably lower than the characteristic impedance and the load end of the line must be matched to an unusually very small impedance level (the characteristic impedance of the line). For the degenerated CX delay, the usage of these buffers is not so relevant. It is usually possible to reduce loading effects using appropriate design strategies. Furthermore, the necessity of very low impedances or matched conditions is not mandatory for this case.

The implementation of both active and transmission line delay elements using current integrated circuit design process is dramatically affected by power and area constraints. This thesis advocates the usage of delay elements in feedback amplifiers as a means of achieving significant bandwidth improvements. Feedback amplifiers are usually part of analog processing circuitry in large scale integrated circuits, occupying small silicon area and obeying stringent power restrictions. Transmission lines usually require large design areas (due to current density restrictions and coupling reduction) and large power (small characteristic impedances imply large currents even for small voltage swings). From this perspective, active delay elements are especially suited for the integrated circuit environment, since they can be designed in order to meet both low power and low area constraints.

Another important aspect is the controllability of the delay. It is often necessary to provide external control over some design variable in order to cope with process variations or for design optimization. One such example is the case of bandwidth optimization in delayed feedback amplifiers. Active delay topologies can provide easy to implement delay control mechanisms (for instance controlling the delay time constant). Transmission lines, on the other hand, have fixed nature implying fixed delays and are especially susceptible to process variations.

From the above discussion it can be concluded that active delays can satisfy all the common

integrated circuit design constraints, and also provide easy to implement control mechanisms, making them the preferred choice. Nevertheless, transmission lines become a preferable solution for high frequency applications. As frequency increases, the necessary delays decrease thus reducing the area penalties: several recent contributions report the usage of small transmission lines in CMOS transceiver applications for the 60 GHz band [192-194]. Furthermore, special purpose integrated circuits using high performance integration technologies like GaAs, InP or SiGe may also consider the usage of transmission lines as delay elements, since in general area and power penalties are of no concern for these special purpose applications.

4.5 Transimpedance Amplifier Design Considerations

Transimpedance amplifiers are perhaps the most important of all the feedback amplifier configurations. Their widespread usage is in part justified by their adequacy to the conversion between current and voltage mode circuitry. This is also the main reason for their application to modern optical communication systems (see chapter 1 for a brief exploitation of this statement). In such applications, it is commonly necessary to convert a current resulting from an optical detection device into a voltage signal suitable for further processing. Since optical detection devices are mainly PIN (P type – Intrinsic – N type) or APD (Avalanche Photodiodes) photodiodes, several challenging design considerations must be satisfied, namely:

- i) The optical-electrical conversion usually results in small currents (also dependent on the medium and distance between emitter and receiving equipment), that need high gain amplification factors;
- ii) Photodiodes are generally modeled as a current source with some parallel capacitance, which poses severe bandwidth limitations (especially for free-space optical communication systems where the required photodiodes have large exposure areas and consequently high intrinsic capacitance);
- iii) Noise is also an important factor that can compromise the decision circuitry. Noise minimization usually requires large impedance biased photodiodes, which results in bandwidth limitations (chapter 1 presented a simplified description of these effects);
- iv) DC level fluctuations. The signal detected by the photodiode is only a component of the total optical-electrical converted current. A great percentage of this current is intimately related to the light exposure of the photodiode, once again this is especially relevant for the free-space optical case (this is perfectly controlled in fiber optic communication systems).

There are several techniques to combat these challenging problems namely:

- i) Using high gain transimpedance amplifiers able to amplify and convert the input current into

- an output voltage, suitable for processing;
- ii) Transimpedance amplifiers have small controllable input impedances suited to amplify current signals from capacitive sources without posing severe limitations on bandwidth;
- iii) Noise minimization is generally accomplished using adequate noise matching strategies at the input of the amplifier; and
- iv) DC level fluctuations can be filtered using optical and/or electrical filtering (integrated circuit techniques often recur to DC current sensing and subtraction mechanisms in order to avoid the usage of high DC blocking capacitances).

High gain and high bandwidth are perhaps, the most challenging combination between these problems. The gain bandwidth product is fixed by technological issues as discussed in Chapter 2. This means that gain and bandwidth have inverse dependencies; increasing one reduces the other and vice-versa. Frequency peaking strategies are usually employed to combat this inverse dependency, allowing an effective exploitation of the maximum allowed bandwidth for a prescribed gain level. Delayed feedback in this respect is not different from other frequency peaking strategies. However, due to its simplicity, it is easily extendable to all feedback topologies (this is not always true for other frequency peaking strategies that may rely on certain preferred topologies). Transimpedance amplifiers can be bandwidth optimized using the concept of delay feedback together with the designed procedure presented on chapter 3 and the results of chapter 4. The following section discusses transimpedance amplifier design problems, using delayed feedback optimization.

4.5.1 Closed-Loop Gain and Return Ratio

Figure 4.26 depicts a general transimpedance amplifier comprising one delay element inside its amplifying chain. In order to reflect optical reception framework, a current source comprising internal conductance G_S and capacitance C_S were used. Two internal buffers A_1 and A_2 are employed due to delay element matching considerations (as discussed on section 4.3.5). The delay element is represented by its equivalent y parameter representation $[y]_D$ (as discussed on sections

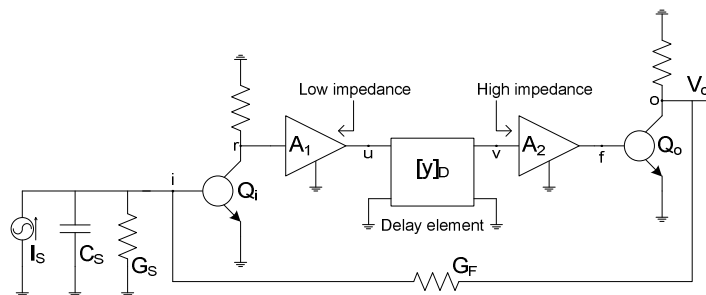


Figure 4.26 – Transimpedance amplifier small signal equivalent.

4.4.2 and 4.4.3). Two generic transistors Q_i and Q_o represent the amplifier's input and output ports. Due to signal amplitudes and noise considerations the stage with larger gain must be placed at the beginning of the amplifying chain [270, 271, 279, 280, 284]; this is accomplished using CX or Cascode topologies. For the present case, a simple CX stage, consisting on Q_i was considered. The output stage must have small output impedance in order to reduce feedback loading effects at the output port. For the present case a CX stage was also chosen for the output stage. This is not the best choice; a CZ stage is more adequate to fulfill the low output impedance requirement. However, the choice between these topologies is conditioned by the number of phase reversals between the input and output of the amplifying chain, which must be an odd number in order to have negative feedback. For a well designed (stable) transimpedance amplifier the number of amplifying stages should be the least possible; since for each new added stage there will be at least one more pole and hence reduced stability margins. Since the delay element can comprise phase reversing stages (A_1 , A_2 or the delay itself, if an active delay topology is in use), the last stage should be used to provide the necessary phase condition and also reduce the output impedance to an acceptable level. This is also the main reason to include the delay on the forward path. As it was discussed in chapter 3, the presence of a delay element inside the feedback loop is not restricted to any particular branch. However, from a circuit design perspective, placing the delay on the forward branch of a feedback amplifier represents a possible economy on gain stages. In fact, it is possible to accommodate adequately all the necessary phase reversals, gain conditions, and stability requirements, using the matching buffers. Finally, the feedback network is represented by the conductance G_F between input and output terminals. The input and output stages were represented by CX configurations with resistive load in order to ease the forgoing analysis. Nevertheless, it is possible to consider active load topologies often used in integrated circuit design.

The closed-loop gain and the return ratio can be evaluated using the asymptotic gain model presented on section 4.2. Considering a detailed small equivalent circuit represented on Figure 4.27: the input and output stages are represented by their equivalent y parameter sets, where only the relevant part is depicted (y_i , y_r representing the input port due to Q_i and y_f , y_o representing the output port due to Q_o); the fractional gain H_i^r represents the voltage gain of the input stage and defines the voltage v_r (controlling the internal feedback of Q_i); the voltage gain from between the input terminal and the Y terminal of Q_o (node f) is $H_i^r A_1 D A_2$, where A_1 and A_2 are the voltage gain due to buffers

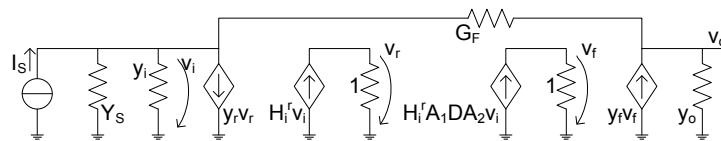


Figure 4.27 – Detailed small signal equivalent of the circuit of Figure 4.26.

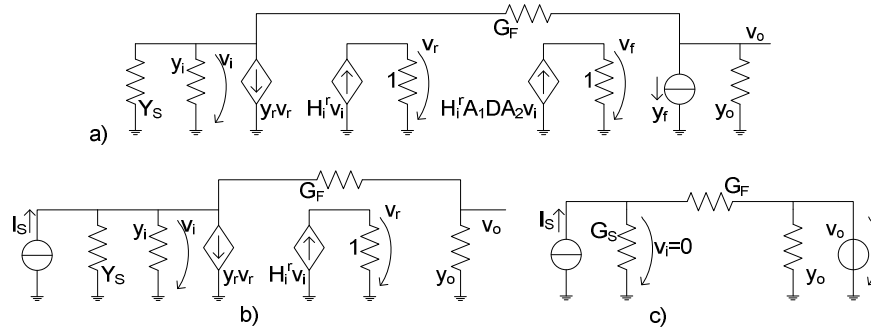


Figure 4.28 – Circuit modifications to evaluate: a) return ratio; b) direct gain; c) asymptotic gain.

A_1 and A_2 and D is the voltage transfer function of the delay element (obviously the aforementioned gain factors include loading effects; finally Y_S represents the source admittance (comprising conductance G_S and capacitance C_S). The return ratio can be defined in order to any controlled source in the circuit, using the controlled current source of the Q_o (y_f), the return ratio is evaluated assuming the modifications represented on Figure 4.28 a), resulting in,

$$R = -V_f = \frac{y_f G_F H_i^r A_1 D A_2}{(y_o + G_F)(y_i + Y_S + G_F + y_r H_i^r) - G_F^2} \quad (4.58)$$

The frequency dependence of (4.58) is implicitly represented by the admittance terms and the voltage gain transfer functions. The direct gain term is the remaining transfer function when the reference controlled source is set to zero, that is, using the simplified circuit of Figure 4.28b,

$$G_0 = \lim_{y_f \rightarrow 0} \left(\frac{V_o}{I_s} \right) = \frac{G_F}{(y_o + G_F)(y_i + Y_S + G_F + y_r H_i^r) - G_F^2} \quad (4.59)$$

Finally, the asymptotic gain is the transfer function between input (I_S) and output (V_o) when the reference controlled source becomes infinite. When y_f goes to infinity, the current $y_f V_f$ must remain finite. The unique possibility is to set V_f to zero. Since V_f has a linear dependence on V_i this is the same as forcing V_i to zero. Using the equivalent circuit of Figure 4.28 c), the asymptotic gain term is found by the relation between the auxiliary voltage source V_o at the output port and the input current source I_S , when V_i is equal to zero,

$$G_\infty = \lim_{y_f \rightarrow \infty} \left(\frac{V_o}{I_s} \right) = -\frac{1}{G_F} \quad (4.60)$$

Using other controlled source as reference produces different values for all the previous entities. Nevertheless, the closed-loop transfer function results the same, as long as the input and output variables are preserved. The closed loop transfer function is written as,

$$H(s) = \frac{G_\infty R(s)}{1 + R(s)} + \frac{G_0(s)}{1 + R(s)} \quad (4.61)$$

Since $R(s)$ appears in both terms of (4.61), the delayed feedback concept applies to both asymptotic and direct terms. In general the direct term $G_o(s)$ has negligible effect on (4.61) facing $G_\infty R(s)$. For these cases is sufficiently accurate to consider only the first term. Whenever such simplifying assumption fails, the complete form of (4.61) must be considered. Comparing the simplified version of (4.61) with the delayed feedback formulation, the magnitudes of the loop gain G_o , and the forward gain A_o are represented by $|R(0)|$ and $|G_\infty R(0)|$ respectively. Accordingly, maximum flatness and stability constraints are expressed by $|1+R(j\omega)|^2 > |G_\infty R(j\omega)|^2$ and $1+R(s)=0$, respectively.

4.5.2 Pole and Zero Contributions

According to (4.61) the open-loop poles and zeros are respectively the singularities and the zeros of $R(s)$. Once again the superiority of the asymptotic gain model for design shows its evidence; the open-loop dynamic parameters are all expressed by the return ratio²⁹. The return ratio expression can be rearranged in order to disclose all the amplifying stages contributions, and thus,

$$R = \frac{H_i^r}{y_i + Y_s + G_F + y_r H_i^r - \frac{G_F^2}{y_o + G_F}} \frac{y_f G_F A_1 D A_2}{y_o + G_F} \quad (4.62)$$

The first term in (4.62) represents the input circuitry dependence, including loading effects due to the feedback through G_F . The second term represents the rest of the amplifying chain: matching amplifiers A_1 and A_2 , delay element D and output stage plus loading.

Assuming as before that the first amplifying stage provides the highest gain factor, and consequently sets the bandwidth limitations of the overall cascade and that the subsequent stages have negligible frequency dependence (except for the delay elements that may be assumed ideal), only the first term in (4.62) is significant from the frequency response point of view. A close inspection also reveals that this term has impedance dimensions, thus it can be adequately represented as a first stage transimpedance gain factor $Z_i^T(s)$. Using the y parameter representation of the CX input stage, the factors y_i , y_r and H_i^r are suitably expressed as functions of the small circuit equivalent on Figure 4.29,

²⁹ It is important to notice that this is a consequence of the assumed reference controlled source. Using another reference may lead to different frequency dependent terms; for instance, using G_F as the controlling element (using an y parameter representation and taking the reverse transmission as reference), results in $G_\infty=0$. For this example, the second term of (4.61) is itself the complete closed-loop transfer function.

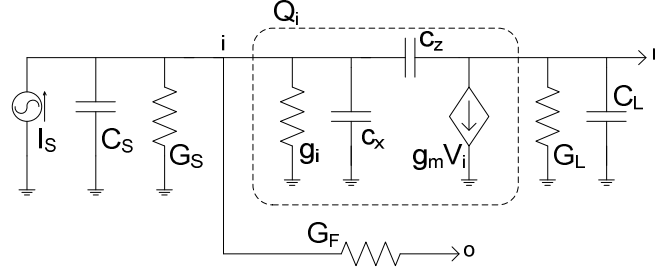


Figure 4.29 – Small signal input equivalent circuit of the amplifier in Figure 4.26.

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} g_i + s(c_x + c_z) & -sc_z \\ g_m - sc_z & G_L + s(c_z + C_L) \end{bmatrix} \quad (4.63)$$

where G_L and C_L represent the load of the first stage. Now, equating y_i , y_r and H_i^r results,

$$\begin{aligned} y_i(s) &= y_{11}(s) = g_i + s(c_x + c_z) \\ y_r(s) &= y_{12}(s) = -sc_z \end{aligned} \quad (4.64)$$

and,

$$H_i^r(s) = -\frac{y_{21}(s)}{y_{22}(s)} = -\frac{g_m}{G_L} \frac{1 - sc_z/g_m}{1 + s(c_z + C_L)/G_L} \quad (4.65)$$

Substituting (4.64) and (4.65) into the first term of (4.62) reveals that the return ratio frequency dependent term due to the input circuitry is expressed by,

$$Z_i^T(s) = -Z_i^T(0) \frac{1 - s/z}{1 + \frac{2\xi s}{\omega_n} + \left(\frac{s}{\omega_n}\right)^2} \quad (4.66)$$

with $Z_i^T(0)$, z , ξ , and ω_n defined by,

$$\begin{aligned} Z_i^T(0) &= \frac{g_m/G_L}{g_i + G_s + \frac{y_o G_F}{y_o + G_F}} \\ z &= g_m/c_z \\ \omega_n &= \sqrt{\frac{g_m/Z_i^T(0)}{c_z C_L + (c_z + C_L)(c_x + C_s)}} \\ \frac{2\xi}{\omega_n} &= Z_i^T(0) \frac{G_L}{g_m} \left(C_s + c_x + c_z \left(1 + \frac{g_m}{G_L} \right) \right) + \frac{c_z + C_L}{G_L} \end{aligned} \quad (4.67)$$

An alternative representation shows that the second order denominator in (4.66) results from the interaction of two poles with different frequencies, where the term of the first power of s represents the sum of the time constants associated with the two poles (also known as first order time constants). The frequencies associated with these first order time constants are defined by,

$$p_1 = \frac{g_i + G_S + y_o G_F / (y_o + G_F)}{C_S + c_x + c_z (1 + g_m / G_L)} \quad (4.68)$$

$$p_2 = \frac{G_L}{c_z + C_L}$$

Since the second order time constant (the reciprocal of ω_h^2) is not the product of the reciprocals of p_1 and p_2 , the open-loop poles are badly approximated by (4.68). It is possible to have a closer approximation of the open-loop poles, defining the pole multiplying factor as,

$$M = \left[1 + \frac{c_z}{c_x + C_S} (1 + g_m / G_L) \right] \left[1 + \frac{c_z C_L}{(c_x + C_S)(c_z + C_L)} \right]^{-1} \quad (4.69)$$

Then using (4.68), (4.69) and assuming that $p_2 \gg p_1$, the open-loop poles are approximately given by

$$p_1^* \approx \frac{M}{2} p_1 \quad , \quad p_2^* \approx M p_2 + \frac{M}{2} p_1 \quad (4.70)$$

p_1^* is usually the dominant pole, due to the presence of the source capacitance C_S . This is particularly true for free space optical (FSO) receiving systems. In an optical receiving front-end, C_S represents the intrinsic capacitance of the photo-detector. There are two relevant design scenarios: i) fiber optic systems (FO) – where the photo-detectors have small intrinsic capacitance; and ii) FSO systems – where, due to efficiency restrictions the photo-detectors have large photosensitive areas and consequently, large intrinsic capacitances.

Equations (4.67), (4.68), (4.69) and (4.70) reveal several important observations:

- i) The dominant pole contribution is especially affected by the source capacitance C_S and the Miller multiplied capacitance c_z ;
- ii) The conductance factor in p_1^* is dominated by the first stage transistor input conductance g_i , and by the conductance due to feedback network plus loading. The source conductance (G_S) is normally lower than the aforementioned conductances and thus can be neglected;
- iii) Since large closed-loop transimpedance gains require even larger open-loop transimpedance gains, this is best accomplished using a first stage having both high input resistance and voltage gain;
- iv) It is also important, due to stability considerations, to have a large separation between the first and second poles. Since the second pole, p_2^* , is essentially dominated by p_2 , this is accomplished using large conductances for G_L and lowering C_L (note that C_L is essentially formed by parasitic and transistor intrinsic capacitances);
- v) Gain and pole separation are traded by bandwidth. p_1^* decreases, when the first stage transimpedance gain is increased. Increasing the pole separation implies low values of G_L

and consequently, high Miller effect on the first stage.

The first stage transimpedance gain can be evaluated from the required closed-loop gain. Since the asymptotic gain is known, assuming that the direct gain term is negligible compared to $R(0)$,

$$Z_i^T(0) = -\frac{H(0)}{H(0) - G_\infty} \frac{y_o + G_F}{y_f G_F A_1 A_2 D(0)} \quad (4.71)$$

Once the voltage gain of all the amplifying stages is set, equation (4.71) can be used to compute the necessary first stage transimpedance gain. Equations (4.67) and (4.68) reveal that p_1 and $Z_i^T(0)$ have a close relation. In fact, it is possible to write,

$$Z_i^T(0) p_1 = \frac{g_m / G_L}{C_s + c_x + c_z (1 + g_m / G_L)} \quad (4.72)$$

According to (4.72), the product of the first stage's transimpedance gain by the input pole (dominant pole) is fixed by the ratio of the first stage's voltage gain by the total equivalent input capacitance of the stage. This relation shows that the dominant pole is generally a function of the required gain for this type of amplifying configurations. Since $Z_i^T(0)$ is ultimately set by the required closed-loop gain, according to (4.71), this means that, changing the feedback conductance G_F also impairs the dominant pole of the circuit. This feedback dependence is particularly evident in transimpedance amplifiers and implies a necessary transformation of the procedures for bandwidth evaluation discussed in chapter 3, section 3.5.1. This will be discussed in more detail in chapter 6, using a transimpedance amplifier design example as reference.

4.5.3 Delay Contributions

The previous section discussed a modeling approach suitable for the design of a delayed feedback amplifier. To finish this discussion there is one final design parameter to consider, the total loop delay. As it was presented on section 4.1 the total loop delay consists of two main contributions: intrinsic delays due to transistors and the intentionally added extrinsic delays. Section 4.3 discussed the nature of intrinsic delays for various transistor types. Section 4.4, on the other hand, presented two methods suitable for the design of extrinsic delays. Extrinsic delays can be further classified according to the design method. Thus, extrinsic delays can be of all-pass type (consisting on suitable implementation of Padé approximants of the complex exponential), or of transmission line type (consisting on a meander transmission line). There is also another source of delay to consider associated with the presence of non-minimum phase terms in circuit transfer functions. As it was discussed in 4.4.1 and 4.4.3, these non-minimum phase terms consist of RHP zeros which can be described as Padé approximants of the complex exponential.

Once these delay sources have been properly identified, the total delay is found using the

following equation,

$$L_{eff} = \sum_{Intrinsic} \tau_i + \sum_{All-Pass} T_i + \sum_{TL} L_i + \sum_{RHP} z_i \quad (4.73)$$

In (4.73), each summation is taken in respect to the total delay elements having the same nature, that is, τ_i for all the intrinsic delay contributions, T_i all the all-pass delay contributions, L_i for all complex exponential delays and z_i for all RHP zeros.

4.5.3.1 Intrinsic Delays in Electronic Circuits

Section 4.3.2 showed that intrinsic delays are usually modeled as complex exponential terms appended to the controlled current source of the small signal equivalent model of the transistor. There are three equivalent transistor representations using different controlled current sources: i) the typical Π -model uses a voltage controlled current source, with the transistor's transconductance g_m as control parameter; ii) the T-model, also known as common base model (or common gate), which uses a current controlled current source, with the transistor α current gain (between Z and X terminals) as control parameter; and iii) the hybrid-model which uses a current controlled current source, with the transistor β current gain (between Z and Y terminals) as control parameter. These control parameters are related using standard conversion formulas.

$$\begin{aligned} \alpha &= \frac{\beta}{\beta + 1} \\ \beta &= g_m r_i \end{aligned} \quad (4.74)$$

The effect of the appended delay on each of these three control parameters is strongly dependent on two factors: the selected equivalent model and the circuit configuration. For instance, from (4.74) it is readily seen that the same delay must apply for β and g_m , however, this is not necessarily the same for α . Equation (4.18), on section 4.3.1 shows an approximate delay conversion between α and β modeling strategies.

Similar problems arise due to circuit configurations. Taking the Π -model as reference, the transconductance parameter may not appear only as a gain proportionality factor. Certain configurations exhibit both direct and inverse proportionality dependencies on g_m . This is a general consequence of circuit configurations using local feedback. For instance, degenerated CX, and CZ stages exhibit this phenomenon, since the degenerating/load resistance forms a feedback loop between the input and output ports. These cases can be treated using the delayed feedback formulation. However, due to their simplicity and widespread usage it is preferable to consider approximation procedures. Three such approximations are (ordered by their increasing accuracy): i) to neglect the delay effect – this approach is suitable for circuits employing high frequency

transistors operating in a medium frequency range; ii) convert the complex delay element appearing on the denominator of the transfer function using a first order Taylor representation – this approach can impair the gain accuracy, nevertheless it improves the phase characteristics; and iii) convert the complex delay element appearing on the denominator of the transfer function using a first order numerator-denominator Padé representation – this is the most accurate and simple approximation of the three.

4.5.3.2 RHP Zeros on Electronic Circuits

RHP zeros are usually related to circuits exhibiting Miller capacitance multiplication, such as CX or degenerated CX stages. On section 4.3.3 this property was explored as a means of synthesizing simple and accurate Padé approximants based on degenerated CX stages. However there are many other circuits on which RHP may manifest. For instance, the subtraction of two low-pass transfer functions can originate one such RHP zero³⁰,

$$\frac{k_1}{1+s\tau_1} - \frac{k_2}{1+s\tau_2} = \frac{(k_1-k_2)}{(1+s\tau_1)(1+s\tau_2)} \left(1 - s \frac{k_2\tau_1 - k_1\tau_2}{k_1 - k_2} \right) \quad (4.75)$$

When the RHP zero appears in a Padé approximant form (as discussed in section 4.3.3), the delay contribution is clearly acknowledged. However, if such zeros appear in a isolated fashion, the overall transfer function has to be converted into the product of a minimum phase transfer function by a suitable number of Padé approximants, for each RHP zero, that is,

$$H(s) \prod_{RHP}^n (1 - s/z_i) = \left[H(s) \prod_{LHP}^n (1 + s/z_i) \right] \prod_{Padé}^n \frac{(1 - s/z_i)}{(1 + s/z_i)} \quad (4.76)$$

The delay arising due to these RHP zeros is now clearly evaluated as,

$$\sum_{RHP} T_i = \sum_1^n \frac{2}{z_i} \quad (4.77)$$

The extra LHP zeros on (4.76) can be neglected if their combined time constants (half the value of the time delay in (4.77)) is negligible when compared with the circuit's dominant time constants.

4.6 Concluding Remarks

This chapter discussed the theoretical aspects of the design of delayed feedback amplifiers. Section 4.2 introduced the asymptotic gain model as the most suited analytic method for the analysis of feedback amplifiers. Sections 4.3 and 4.4 discussed the nature of delays in electronic

³⁰ This property may also serve the purpose of synthesizing Padé approximants in current-mode circuits, using $k_I=2$, $k_I=1$ and $\tau_2=0$.

circuits and ways to design delay elements. Two general types of delays can be considered, intrinsic and extrinsic delays. Intrinsic delays are generally due to electronic devices that make part of the amplifier circuit. Section 4.3 discussed intrinsic delays due to transistors. Both bipolar and FET devices were considered due to their different current conduction mechanisms. Section 4.4 discussed two kinds of extrinsic delays, transmission line delays and all-pass type delays. Both of these delay kinds are suitable for implementation using appropriate circuit techniques. Finally, section 4.5 explored the dynamics of a transimpedance amplifier employing delayed feedback methods.

5 CURRENT-MODE TECHNIQUES

“You’re driving alone at night. And it’s dark and it’s raining. And you took a turn back there and you’re not sure now that it was the right turn, but you took the turn anyway and you just keep going in this direction. Eventually, it starts to get light and you look out and you realize you have absolutely no idea where you are.” – Laurie Anderson.

Summary

This chapter addresses the usage of current-mode design techniques as a means of achieving large GBW products in transimpedance amplifiers for large capacitive sources. Current-mode design techniques are known by their inherent ability to convey signals in highly capacitive mediums without posing severe limitations on bandwidth. Following this line of reasoning, this chapter introduces the concept of active current-matching. According to this concept a CMD (current matching device) is placed between the capacitive source and the transimpedance amplifier’s input. The GBW limitations can be reduced if the input impedance of the CMD is smaller than the input impedance of the transimpedance amplifier.

5.1 Current-Mode Design Techniques

Modern processor design industry has been the major driving force dictating the constant and fast evolution of integrated circuit technology, especially CMOS processes. Since the early seventies, after Moore [213, 220, 242] and Dennard [217-219, 221] studies, that the benefits of scaling down the size of MOS transistors are known. Moore observed that integrated technology increased the integration capabilities with a rate of roughly twice a year since the early sixties [213]. Dennard introduced the principles of scaling, according to which faster and smaller transistors would be possible [217]. Since then, Dennard scaling theory has been “Moore’s law” sustaining principle. This rapid evolution brought to circuit design important changes. One of these changes was the advent of bias voltage reduction. The first scaling strategies used the well known constant field scaling principles consisting in the simultaneous reduction of all the voltages and dimensions by the same factor α , which improves circuit speed and density by a factor of α and α^2 respectively [217-219]. Constant field scaling predicts ever decreasing bias and threshold voltages; although reality has not been as straight as predicted. The principal reason is that the thermal gap (kT/q) is one quantity that can not be scaled. Generalized scaling and constant voltage scaling strategies have been proposed to cope with this problem [221]. Nevertheless, bias voltage has reached very small values (below 1V), which are not well suited for analog design.

Designing analog electronic circuits able to operate with low bias voltages is by itself a challenging problem. Adding to this the fact that, for every technological node jump, bias voltages are getting closer to the threshold voltages, turns analog design issues into a real nightmare! Digital circuit design is not as dramatically affected by these voltage constraints, since transistors are normally used as switches. The same is not true in analog circuits. Analog circuits, often require large linear operating ranges which are difficult to realize with these voltage constraints. The reasons for these difficulties are mainly due to two facts: i) first, small bias voltages pose a limit on the maximum possible output voltage swing; ii) second, transistors with threshold voltages comparable to the bias are highly non-linear devices. Techniques able to cope with these voltage constraints have been published [259, 263, 265, 276, 278, 281]. These techniques aim to extend the linearity of the active devices and also to provide an efficient usage of all the available bias.

Voltage constraints are particularly important in the so called voltage-mode circuits, that is, circuits where information signals are represented by node voltages, mainly because all the circuit node voltages have a direct dependence on the bias voltage. An obvious alternative to voltage-mode circuits are current-mode circuits, that is, circuits where information signals are represented on the current domain [263, 265, 276, 278, 281]. This is also a natural choice since transistors are current

controlled devices: the current between drain and source in FET transistors is controlled by the voltage applied from gate to source; the current between collector and emitter in bipolar transistors depends on the current injected in the base terminal. The major advantage of current-mode circuits lies on the fact that the current span is not directly related to the bias voltage, making current-mode circuits specially suited for operation under low bias voltages. Considering the voltage compression characteristics of the transistor (logarithmic laws for bipolar transistors and square root laws for FET), it is a simple matter to show that large current spans correspond to compressed voltage swings. Nevertheless, equal voltage and current swings can occur whenever a linear current to voltage conversion is implied. These linear current-voltage conversions are seen as something to avoid in current-mode circuits, since bandwidth restrictions are also posed at these interfacing points. Current-mode circuits have also broader frequency responses [259, 276]. Due to the predominantly capacitive nature of integrated circuit's environment, large voltage swings imply large amounts of time to charge and discharge parasitic capacitances. Since current-mode techniques minimize node voltage swings, there will be less time wasted in charging and discharging parasitic capacitances.

Despite all these advantages over voltage-mode circuits, current-mode circuit design techniques have never surpassed their voltage-mode counterparts. The first reported true current-mode processing circuit was the first generation current conveyor (CC-I), proposed by Sedra and Smith in 1968 [215]. They also issued in 1970 an improved version, the second generation current conveyor (CC-II) [216]. However, none of these new elements became an alternative to the traditional operational amplifier and voltage-mode circuits. By that time, voltage-mode techniques based on the operational amplifiers were perfectly established. The concept of operational amplifier was by then very old (around 20 years) [115]. The integrated technology from the sixties and seventies was also at an early stage, and did not provide neither the elements nor the precision required for the implementation of usable current conveyors. It was only during the eighties, when integrated circuit technology introduced the vertical PNP transistor [223, 259, 276], that current conveyors re-earned their deserved attention. Low-voltage operation start revealing the problems and inadequacies of operational amplifiers, which led researchers from industry and academia to devote more attention to current conveyor solutions. Second generation current-conveyors were recognized as universal elements, like the operational amplifier had been many years before [259]. It was possible to synthesize any required transfer function with CC-IIs instead of more traditional operational amplifiers.

Nowadays, current-mode techniques have found applications in many fields. For instance, digital circuit design techniques like ECL (Emitter Coupled Logic) and CML (Current Mode Logic) are

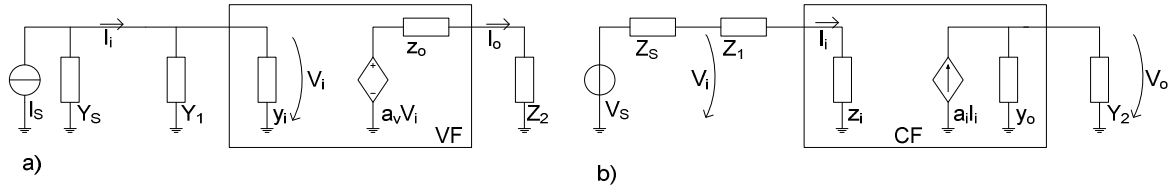


Figure 5.1 - Current (a) and voltage (b) amplification, using VF and CF.

true current-mode design techniques [263]; switched-current circuits are the current-mode version of switched-capacitor circuits [259, 263, 276]; log-domain and translinear circuits are other examples of current-mode applications to signal processing [259, 265, 276].

5.1.1 Current-Mode Design Preliminaries

Chapter 2 introduced the basic ideas behind current-mode design techniques. It was then shown that two important conceptual circuits must be available, a current follower (CF) and a voltage follower (VF) (see Figure 2.33 in section 2.8.4). By definition, a CF is a unitary gain current amplifier having ideally zero input and infinite output impedances. Similarly, a VF has unitary voltage gain, infinite input and zero output impedances. From a bandwidth point of view, these circuits are the least restrictive since the unitary gain condition allows the best possible bandwidth to be achieved. It is possible to synthesize any possible input/output relation using these two circuits and some pure resistances. For this situation, the achieved bandwidth is solely restricted by the bandwidth of the voltage or current buffers in use. However, practical circuits can only approach this ideal behavior. The presence of non-ideal input and output impedances poses another kind of bandwidth restrictions, due to current or voltage division at these ports. These limitations can be adequately appreciated taking Figure 5.1 as reference.

Figure 5.1a represents the current-mode approach to current amplification. It is easily seen that for the ideal case, where y_i is infinite and z_o zero, the current gain (taking I_i as reference for the input) is set by the reciprocal of the product of Z_2 times Y_1 , multiplied by the voltage gain of the VF (ideally unitary gain). Considering that the VF has frequency dependent terminal impedances and voltage transfer, the current gain can be written as:

$$A_i(s) = \frac{I_o(s)}{I_i(s)} = \frac{a_v(s)}{Y_1(s)Z_2(s)} E_i(s) E_o(s) \quad (5.1)$$

where $a_v(s)$, $Y_1(s)$ and $Z_2(s)$ represent the frequency dependent voltage gain and auxiliary impedances for gain establishment. The factors $E_i(s)$ and $E_o(s)$ quantify the frequency dependent error due to current division in $y_i(s)$ at the input, and voltage division due to $z_o(s)$ at the output, which are given as:

$$E_i(s) = \frac{Y_1(s)}{Y_1(s) + y_i(s)} \quad E_o(s) = \frac{Z_2(s)}{Z_2(s) + z_o(s)} \quad (5.2)$$

Assuming that both $a_v(s)$, $y_i(s)$ and $z_o(s)$ can be approximated by first order models (for the cases of $y_i(s)$ and $z_o(s)$, this corresponds to a simple parallel association of one resistance and one capacitance, thus $y_i(s) = g_i + sC_i$ and $z_o(s) = r_o / (1 + s r_o C_o)$), and that $Y_1(s)$ and $Z_2(s)$ consist off a pure resistance (G_1 and R_2 , respectively) and some parallel parasitic capacitance (C_1 and C_2 respectively), the complete transfer function becomes,

$$A_i(s) = \frac{a_v(0)}{(G_1 + g_i)(R_2 + r_o)} \frac{(1 + s\tau_{o\infty})(1 + s\tau_2)}{(1 + s\tau_v)(1 + s\tau_{iL})(1 + s\tau_{oL})} \quad (5.3)$$

where τ_v , τ_{iL} , τ_{oL} , τ_2 , $\tau_{o\infty}$, are circuit time constants for:

- τ_v – is the voltage transfer function time constant;
- τ_{iL} – is the time constant due to the input port loaded by Y_1 (given by $(C_1 + c_i)/(G_1 + g_i)$);
- τ_{oL} – is the time constant due to the output port loaded by Z_2 (given by $(R_2/r_o)(C_2 + c_o)$);
- τ_2 – is the time constant due to the load (given by $R_2 C_2$); and
- $\tau_{o\infty}$ – is the time constant due to the output without load (given by $r_o C_o$).

At a first glance it seems that the non-idealities of the VF act in conjunction with Y_1 and Z_2 to pose limitations on the achievable bandwidth of the final current amplifier. However this is not necessarily true. Usually the major bandwidth restrictions are due to the input circuitry (τ_{iL}) and VF (τ_v), so the major design concern should be to cancel these effects. Since $\tau_{o\infty}$ is presumably the least significant time constant, it is possible to use pole-zero cancellation, using τ_2 to cancel τ_{iL} . Equating τ_2 to τ_{iL} results in the following design condition,

$$\frac{c_i + C_1}{C_2} = \frac{a_v(0)}{A_i(0)} \frac{R_2}{r_o + R_2} \quad (5.4)$$

According to (5.4), the value of C_2 (including parasitic capacitances) is adjusted in function of C_1 and R_2 once the desired gain, $A_i(0)$ is known. Using this strategy, the pole due to τ_{iL} is cancelled, and substituted by another pole due to τ_{oL} . The effectiveness of this technique can be measure by the ratio between τ_{iL} and τ_{oL} , which is given as,

$$\frac{\tau_{iL}}{\tau_{oL}} = a_v(0) \frac{C_2}{C_2 + c_o} \left(1 + \frac{R_2}{r_o} \right) \quad (5.5)$$

Since C_2 grows proportionally with C_1 and $A_i(0)$, the ratio between τ_{iL} and τ_{oL} is potentially larger than 1 for current gains larger than 1. Using pole-zero compensation results in complexity reduction of (5.3) and also improves bandwidth. Nevertheless it should be noted that the two

remaining poles due to τ_v and τ_{oL} can be rather close, depending on the value of C_2 used to compensate τ_{iL} .

Figure 5.1b presents how voltage amplification is achieved using current-mode techniques. Instead of a VF as in Figure 5.1a voltage amplification needs a CF. Since the circuits of Figure 5.1a and b are a pair of dual circuits, the complete transfer function for the circuit of Figure 5.1b can be derived from (5.1) and (5.2) interchanging impedances by admittances and vice-versa and using $a_i(s)$ instead $a_v(s)$, thus,

$$A_v(s) = \frac{V_o(s)}{V_i(s)} = \frac{a_i(s)}{Z_1(s)Y_2(s)} E_i(s) E_o(s) \quad (5.6)$$

where,

$$\begin{aligned} E_i(s) &= \frac{Z_1(s)}{Z_1(s) + z_i(s)} \\ E_o(s) &= \frac{Y_2(s)}{Y_2(s) + y_o(s)} \end{aligned} \quad (5.7)$$

The observations concerning the frequency behavior made for the case of current amplification are also valid for the present case. The major difference is that now zero contributions arise from the input circuitry rather than from the output, as before. As before, pole-zero compensation is possible using the loaded output pole to compensate the input zero due to R_I and C_I .

Transimpedance and transadmittance amplification require more elaborate attention. Starting with the transimpedance case, it is possible to use either CF or VF strategies to achieve the same transimpedance gain. One possibility is to use the circuit set-up of Figure 5.1a and taking the output variable as the voltage across Z_2 . Inserting this change in (5.1) results on the following transimpedance transfer function.

$$Z_T(s) = \frac{V_o(s)}{I_i(s)} = \frac{a_v(s)}{Y_1(s)} E_i(s) E_o(s) \quad (5.8)$$

where $E_i(s)$ and $E_o(s)$ are given exactly as in (5.2). Making the same considerations as for (5.3), the complete transfer function is written as,

$$Z_T(s) = \frac{a_v(0)}{G_1 + g_i} \frac{R_2}{R_2 + r_o} \frac{(1 + s\tau_{\infty})}{(1 + s\tau_v)(1 + s\tau_{iL})(1 + s\tau_{oL})} \quad (5.9)$$

where all the entities in (5.9) have the same significance as before. There are two important observations concerning (5.9): i) there is no zero due to R_2 - C_2 as before, so it is difficult to apply pole-zero cancellation strategies; ii) since both transimpedance gain and τ_{iL} are inversely dependent on G_I , increasing the gain results in dramatic losses in bandwidth. Another form to achieve

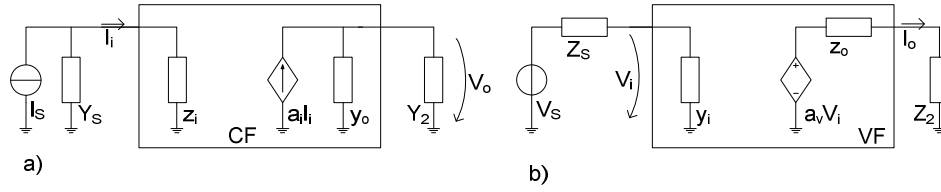


Figure 5.2 – Transimpedance (a) and transadmittance (b) amplification, using CF and VF.

transimpedance gain is to use CF instead of the VF strategy, as depicted on Figure 5.2a. For this case, the only bandwidth limitations arise due to the CF itself and the output circuitry (assuming that the signal source is ideal as before³¹). The transimpedance transfer function for this circuit is written as,

$$Z_T(s) = \frac{V_o(s)}{I_i(s)} = \frac{a_i(s)}{Y_2(s)} \frac{Y_2(s)}{Y_2(s) + y_o(s)} \quad (5.10)$$

For which the frequency response results

$$Z_T(s) = \frac{a_v(0)}{G_2 + g_o} \frac{1}{(1 + s\tau_v)(1 + s\tau_{oL})} \quad (5.11)$$

where $\tau_{oL} = (C_2 + c_o)/(G_2 + g_o)$. The same bandwidth gain duality exists for this case. Nevertheless it can be controlled, since this phenomenon arises here due to the output circuitry.

It is possible to follow the same strategies for the transadmittance case, using CF as in Figure 5.1b, or VF as in Figure 5.2b. Once again, the results of (5.8), (5.9), (5.10) and (5.11), hold the same when replacing impedances by admittances (and vice-versa) and voltage gain by current gain elements.

Although approximate, the above exposition derives important conclusions pertaining the bandwidth limitations in current-mode design, that hold generally true. There are important limitations arising at the input and output circuitry of these current-mode amplifiers, due to the imperfect nature of the available VF and CF. Ultimately, the frequency response is limited by the VF or CF frequency response, if the input-output restrictions are far beyond the cut-off frequency of these elements.

³¹ Considering non-ideal sources as depicted in Figure 5.1 and Figure 5.2, would reveal the frequency behavior of the complete system – amplifier plus source. However this simplified analysis is specially focused on the limitations posed by the amplifier itself without auxiliary circuitry (except for load considerations in some of the cases presented). Nevertheless, including non-ideal sources is a simple matter. It suffices to consider voltage or current division due to the input impedance of the amplifier, which by the way, results in another error factor on all the above expressions.

5.2 Active Current-Matching Concept

Chapter 2 introduced the current-matching concept as a means to break the gain bandwidth duality in transimpedance amplifiers. It is intuitively known, as discussed in chapter 2 and further explored in chapter 4, that gain and bandwidth have inverse dependencies on the desired closed-loop gain in transimpedance amplifiers (and consequently on the feedback resistance). It is also known that the feedback resistance exerts an input loading effect on the main amplifier, causing its open-loop gain to decrease (since the open-loop gain is proportional to the input impedance of the amplifier plus feedback loading). This also has its own effects on the open-loop poles, especially for FSO amplifiers. In such transimpedance amplifiers, the input source is highly capacitive, posing a strong bandwidth restriction on the input port. Techniques able to deal with these capacitive sources often trade gain for bandwidth, at the possible expense of further gain stages prior to the input transimpedance stage (one such case is the well known Cherry-Hopper TAS-TIS configuration [117, 125-127, 129-131]). The current matching concept introduces another design variable, the amplifier input impedance. The next section will show that the input impedance together with the input capacitance form the principal GBW restriction on such transimpedance amplifiers. The closed-loop gain, and hence the feedback resistance exerts only a second order dependence on GBW.

The concept of active current matching device for bandwidth enhancement was originally proposed by Nero Alves in [287]. Prior to this proposal, some authors have used current-mode devices at the input of transimpedance amplifiers with other objectives, for instance in [234, 238] current mode devices were successfully applied to implement gain control and DC bias filtering. Nero Alves *et al* issued several contributions exploring the bandwidth benefits of the current matching concept using CC-II based current buffers [288, 289, 292, 293]. This concept is now attracting researchers attention: in [243, 245] Hou, Godara and Fabre explored the use of current-mode devices as matching buffers between a highly capacitive source and a transimpedance amplifier.

5.2.1 Reference Transimpedance Amplifier

This section will discuss the GBW limitations of FSO transimpedance amplifiers. Figure 5.3 depicts the reference TIA (transimpedance amplifier). In order to make the following discussion more general, it will be assumed that the TIA has second order dynamics; one pole due to the input circuitry and another pole due to the internal amplifying stages (contained in the transimpedance transfer $Z_o(s)$, with gain Z_o and one pole at p_z). The input source is represented by one current source I_p and the capacitance C_p (which may also include the amplifier input capacitance).

Following the asymptotic gain representation of feedback amplifiers presented on chapter 4, the closed-loop transfer function is formed by three auxiliary transfer functions: the asymptotic gain $G_\infty(s)$, the return ratio $R(s)$, and the direct gain $G_o(s)$. These three transfer function are measured

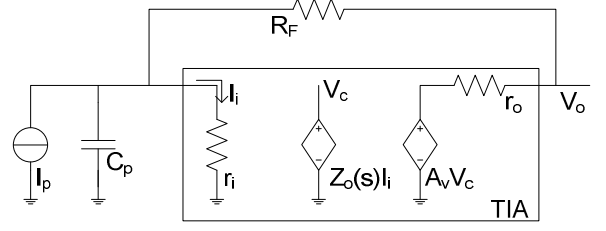


Figure 5.3 – Reference transimpedance amplifier.

against one controlled parameter. For this case the controlled parameter is chosen as the voltage gain, A_v , of the last stage. The return ratio is given as the negative of the control voltage V_c , when the controlled source is replaced by one independent voltage source with value A_v , thus,

$$R(s) = \frac{A_v Z_o}{r_i + r_o + R_F} \frac{1}{(1 + s/p_z)(1 + s/p_i)} \quad (5.12)$$

where p_i is the pole due to the input circuitry, given by $p_i = (r_i + r_o + R_F) / ((r_o + R_F)r_i C_p)$. The direct gain is given as the current to voltage transfer between input and output when the controlled parameter is set to zero. Thus:

$$G_o(s) = \frac{r_o r_i}{r_i + r_o + R_F} \frac{1}{(1 + s/p_i)} \quad (5.13)$$

Finally, the asymptotic gain is found when the controlled parameter goes to infinity. Since, the output voltage must remain finite, this resort to nullifying the control voltage V_c (or equivalently, the current I_i). Then, all the current from the input source must pass through R_F , thus $G_\infty(s) = -R_F$. Using the asymptotic gain formula the overall transimpedance transfer function is given by,

$$\begin{aligned} Z_T(s) &= G_\infty(s) \frac{R(s)}{1 + R(s)} + \frac{G_o(s)}{1 + R(s)} = \\ &\approx -R_F \frac{R(0)}{1 + R(0)} \frac{1}{1 + \frac{p_i + p_z}{(1 + R(0)) p_i p_z} s + \frac{s^2}{(1 + R(0)) p_i p_z}} \end{aligned} \quad (5.14)$$

where the approximation holds whenever $G_o(0) \ll G_\infty R(0)$. Assuming that the amplifier is optimized for maximally flat operation (using appropriate compensation techniques to place p_z adequately, or simply adjusting the gain), the bandwidth of the closed-loop amplifier is given directly by the natural frequency ω_n , and the damping coefficient is $\xi = 1/\sqrt{2}$. Turning to (5.14), this means that,

$$p_i + p_z = 2\xi\omega_n \quad (5.15)$$

and

$$\omega_n^2 = (1 + R(0)) p_i p_z \quad (5.16)$$

Considering that p_z is larger than p_i , and defining d as the pole ratio given by, $d=p_z/p_i$, the GBW becomes,

$$GBW = |Z_T(0)| \omega_n = \sqrt{2} \frac{d}{1+d} \frac{A_v Z_o R_F}{(r_o + R_F) r_i C_p} \quad (5.17)$$

Considering that the two poles have a large separation, that is $d \gg 1$, and that R_F is much larger than r_o , equation (5.17) simplifies into

$$GBW < \sqrt{2} \frac{A_v Z_o}{r_i C_p} \quad (5.18)$$

According to (5.18), the GBW of a transimpedance amplifier having second order dynamics is limited by the ratio between the overall open-loop gain to the input time constant, not considering load effects posed by R_F . Even with R_F , according to (5.17), the loading effects have negligible importance on the achievable GBW. Obviously, the squared root factor comes from the maximally flat design procedure. Higher order dynamics would exhibit larger proportionality factors, according to the discussion on section 2.7.2. Maximizing GBW in transimpedance amplifiers having large capacitive sources is best accomplished reducing the amplifier's input impedance, rather than adjusting gain and/or feedback factor. This justifies the relevance and applicability of the current matching concept.

5.2.2 TIA plus Active Current Matching

As discussed in chapter 2, one way to enhance both gain and bandwidth in transimpedance amplifiers for large capacitive sources, is to detach the source from the input of the amplifier using a current matching buffer in between. This strategy separates effectively the gain from the input impedance, which in common TIA design are intimately related (the open-loop gain is proportional to the input impedance). Figure 5.4 depicts the suggested set-up. An active current buffer having input impedance r_{iC} , output impedance r_{oC} and current gain $a_i(s)$ connects the capacitive source to the transimpedance amplifier. Capacitances C_p and C_i (representing the interstage parasitic capacitance) include all capacitance contributions arising from the CMD (Current Matching Device) and TIA. Since C_p is by assumption, larger than C_i , the gain bandwidth restrictions posed

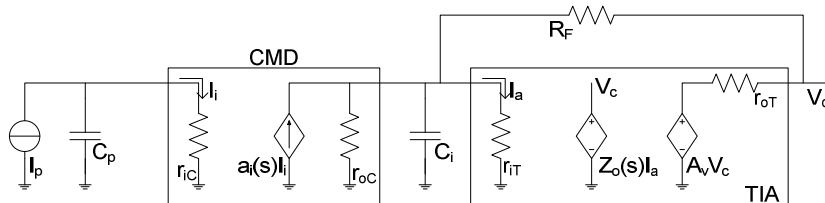


Figure 5.4 – TIA plus CMD design.

on the TIA are relaxed, allowing larger gains for the same reference bandwidth (as in Figure 5.3), or conversely, larger bandwidths for the same reference gain. The previous input bandwidth restriction is now passed to the input of the CMD, which must be designed with low input impedance.

Assuming that $a_i(s)$ is a first order transfer function with gain a_i and one pole at p_a rad/s, the overall transfer function is approximately given by

$$Z_{TC}(s) \approx -\frac{a_i R_F R_C(0)}{1 + R_C(0)} \frac{1}{(1 + s r_{iC} C_p) \left(1 + \frac{s}{p_a}\right) \left(1 + 2 \frac{\xi}{\omega_n} s + \frac{s^2}{\omega_n^2}\right)} \quad (5.19)$$

where,

$$R_C(0) = \frac{Z_o A_v}{r_{oC} // r_{iT} + r_{oT} + R_F} \quad (5.20)$$

and,

$$\begin{aligned} \xi &= \frac{p_{iC} + p_z}{2\omega_n} \\ \omega_n^2 &= (1 + R_C(0)) p_{iC} p_z \end{aligned} \quad (5.21)$$

Equation (5.20) differs from (5.13) due to the fact that now r_i is composed by the parallel association of r_{oC} and r_{iT} . This same factor also has influence on the interstage pole p_{iC} given as

$$p_{iC} = \frac{R_F + r_{oT} + r_{oC} // r_{iT}}{(r_{oC} // r_{iT})(R_F + r_{oT}) C_i} \quad (5.22)$$

One possible way to optimize both gain and bandwidth in (5.19) is to aim for maximal flat frequency response. This is best accomplished if the characteristic equation is represented by a Butterworth polynomial. However, for this case this is not a possible solution. The transfer function in (5.19) has four poles, two real poles and two complex conjugated poles, which does not fulfill the requirements for a 4th order Butterworth polynomial (see appendix A 2). Nevertheless, it is possible to select a set of three poles, one real and two complex conjugated, to form a 3rd order Butterworth polynomial. This is not the best possible solution, nevertheless it allows to draw some interesting theoretical design guidelines to start the optimization process.

Assuming that the pole due to C_p has smaller frequency than the others, it is possible to optimize the remaining three poles in order to synthesize a third order Butterworth polynomial. Starting from the conditions found in section 2.6.2 (equation 2.42), this is accomplished if the inequalities in (5.23) are taken as straight equalities,

$$\begin{aligned} p_a^2 + 2\omega_n^2 (2\xi^2 - 1) &\geq 0 \\ \omega_n^2 + 2p_a^2 (2\xi^2 - 1) &\geq 0 \end{aligned} \quad (5.23)$$

thus implying that $p_a = \omega_n$ and $\xi = 1/2$. Following this strategy requires that CMD must have a current gain cut-off frequency (p_a) equal to the natural frequency of the TIA (ω_n), and that the TIA poles have an associated damping coefficient (ξ) of 0.5. Inserting these conditions on (5.19), the overall frequency response becomes,

$$\left| Z_{TC}(j\omega) \right|^2 \approx - \frac{\left| Z_{TC}(0) \right|^2}{\left(1 + (\omega r_{iC} C_p)^2 \right) \left(1 + \left(\frac{\omega}{\omega_n} \right)^6 \right)} \quad (5.24)$$

The cutoff frequency ω_c , of (5.24) is a function of ω_n and $r_{iC}C_p$. If ω_n is larger than $r_{iC}C_p$, the cutoff frequency is ruled by the input pole and the above considerations are of no practical consequence. However, if ω_n is close to $r_{iC}C_p$, the cutoff frequency is strongly dependent on both factors. Figure 5.5 exhibits the behavior of the normalized cutoff frequency (given as ω_c/ω_n) against the ratio between the input pole ($r_{iC}C_p$) and ω_n .

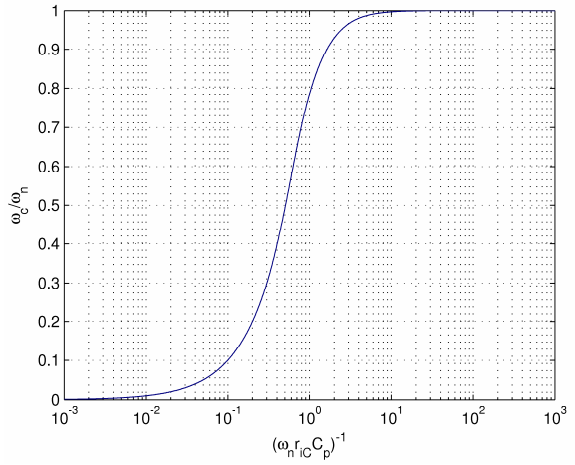


Figure 5.5 – Normalized cutoff frequency as a function of $(\omega_n r_{iC} C_p)^{-1}$.

The GBW achieved by this configuration is determined in a similar fashion as for the case of (5.17). The return difference, $1 + R_C(0)$ is found from the conditions on ξ and ω_n , using (5.21), thus simplifying the gain expression for $Z_{TC}(0)$. Finally the cutoff frequency is expressed as the ratio ω_c/ω_n , resulting in,

$$\begin{aligned} GBW &= Z_{TC}(0) \omega_c \\ &= \frac{a_i A_v Z_o R_F}{(r_{iT} // r_{oC})(r_{oT} + R_F) C_i} \frac{d}{1 + d} \left(\frac{\omega_c}{\omega_n} \right) \end{aligned} \quad (5.25)$$

where d represents the ratio between p_z and p_{iC} .

According to (5.25) the GBW of the TIA+CMD configuration was improved in two aspects: first, the transimpedance GBW is now ruled by the product of the interstage driving point impedance, r_{iT}/r_{oC} by the interstage parasitic capacitance C_i . Since this capacitance is much lower than the source capacitance C_p , the previous bandwidth restriction has been completely removed from the input of the transimpedance amplifier (even for values of r_{iT}/r_{oC} similar to r_i). Second, bandwidth optimization is now accomplished at the input of the CMD stage and thus ruled by the CMD's input impedance r_{iC} . According to (5.25) the value of r_{iC} should be small enough in order to

maximize the ratio ω_c/ω_n . For instance, assuming that $\omega_n r_{iC} C_p = 1$, results in $\omega_c/\omega_n \approx 0.786$; that is, ω_c is just 78.6% of ω_n . $\omega_n r_{iC} C_p$ must be larger than unity in order to have larger ratios of ω_c/ω_n ; equating this in order to r_{iC} , gives

$$r_{iC} < (r_{iT} // r_{oC} // R_F) \frac{C_i}{C_p} \quad (5.26)$$

The inequality (5.26), proves the expected result that the CMD must present low input impedance in order to be useful. (5.26) has another important consequence: since R_F is a direct measure of the asymptotic gain (assuming that both r_{iT} and r_{oC} are larger than R_F and that a_i is close to unity), the maximum value of r_{iC} is given approximately by the product of the asymptotic gain by the ratio of interstage to source capacitances. Thus, larger gains impose less restrictions on r_{iC} than small gains.

It is possible to improve even more the GBW of this configuration if the pair of complex poles in (5.19) is used to add peaking effects near the bandwidth's end. Until now, it was assumed that the three high frequency poles should join to provide a maximum flat Butterworth polynomial, however there is no need to satisfy such constraint. Considering the original set of four poles, it is possible to increase further the bandwidth by simultaneously adjusting the damping coefficient of the pair of complex poles and the position of the other two poles. This can be accomplished using the return difference, since the damping coefficient is inversely proportional to $1+R_C(0)$. This is also a suitable way to enhance bandwidth since it does not impair drastically the overall gain (as discussed previously, the transimpedance gain is proportional to the quotient between the return ratio and the return difference, with the asymptotic gain as proportionality factor – see (5.14)). It is difficult to formally explore this peaking effect since now there is no strict monotonic argument, as for the case of Butterworth polynomials; nevertheless, it is possible to recur to numerical simulation. Figure 5.6

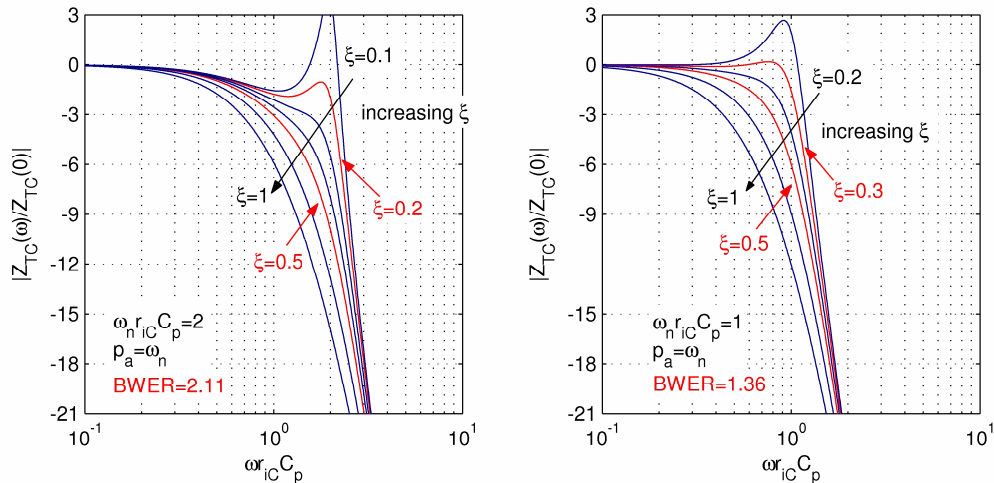


Figure 5.6 – Normalized frequency response, exploring peaking outside the band.

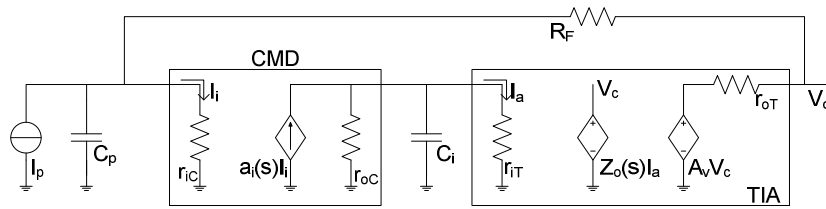


Figure 5.7 – Global feedback TIA plus CMD design.

shows the bandwidth enhancement using this peaking technique for different values of the damping coefficient and two values of $\omega_h r_{iC} C_p$. It is readily apparent that the peaking effect is only effective for small values of $\omega_h r_{iC} C_p$, since large values imply an initial roll-off that can surpass the -3dB reference. Large values of $\omega_h r_{iC} C_p$ require more than one pair of complex poles to promote peaking at other frequencies.

5.2.3 Global Feedback TIA plus CMD

One further possibility to include the CMD on the TIA design is to extend the feedback loop in order to enclose the CMD. Figure 5.7 depicts this global feedback TIA+CMD design strategy; the open-loop transimpedance stage is preceded by one CMD that lowers the open-loop input impedance. Applying feedback through R_F to the overall cascade, results in similar transimpedance gains as before, with reduced values of the input impedance. This strategy has more design freedom than the previous TIA+CMD technique. However, the presence of four real poles does not allow simple optimization schemes as before. The reason for this is that optimizing the frequency response of a closed-loop system with four real poles can be accomplished using several different methods, depending on the magnitude of the poles.

From a stability view point, this is also a worse solution than TIA+CMD. On the previous case, the stability limitations arise due to a 2nd order TIA stage. The remaining two poles are assumed real and hence cannot impair the overall stability. For a global feedback TIA+CMD design, all four poles act within the feedback loop, thus posing stringent stability restrictions.

The closed-loop gain is determined in a similar fashion as before, using the asymptotic gain model. Choosing as before, A_v as the controlled parameter and V_c the control variable, the return ratio is $-V_c$, when I_p is set to zero and the controlled source $A_v V_c$ is substituted by an independent source with value A_v ; that is,

$$R_M(s) = \frac{a_i A_v Z_o r_{oC}}{(r_{iC} + r_{\rho T} + R_F)(r_{\rho C} + r_{iT})} \frac{1}{(1+s/p_i)(1+s/p_a)(1+s/p_{iC})(1+s/p_z)} \quad (5.27)$$

where, p_a the pole due to $a_i(s)$, p_z the pole of the TIA stage, p_i is the pole due to the input circuitry, and p_{iC} the pole at the interstage, which are given by,

$$p_i = \frac{r_{iC} + r_{oT} + R_F}{r_{iC}(r_{oT} + R_F)C_p} \quad (5.28)$$

$$p_{iC} = \frac{r_{iT} + r_{oC}}{r_{iT}r_{oC}C_i}$$

The asymptotic gain is found in a similar fashion as before as $G_\infty(s) = -R_F$. Finally, assuming that the direct gain term has negligible influence on the overall gain, the closed-loop transfer function is approximately given as,

$$Z_{TM}(s) \approx \frac{-R_F R_M(0)}{1 + R_M(0)} \frac{1}{1 + \frac{(1 + s/p_i)(1 + s/p_a)(1 + s/p_{iC})(1 + s/p_z) - 1}{1 + R_M(0)}} \quad (5.29)$$

Equation (5.29) shows that the characteristic polynomial is a 4th order polynomial, consisting of four negative and real open-loop poles. According to the discussion on feedback in chapter 2, this set of open-loop poles cannot provide maximum flatness in strict Butterworth sense. It is difficult to have a precise measure of the maximum bandwidth that can be achieved with (5.29) and the respective closed-loop poles. The natural frequency corresponding to the characteristic equation of (5.29) can be used as an upper bound for the closed-loop bandwidth. Obviously, this upper bound cannot be surpassed.

Thus defining the natural frequency as

$$\omega_n = \sqrt[4]{(1 + R_M(0)) p_i p_{iC} p_a p_z} \quad (5.30)$$

Using (5.30), the GBW upper bound for maximal flat operation is given as

$$GBW < R_F R_M(0) \sqrt[4]{\frac{p_i p_{iC} p_a p_z}{(1 + R_M(0))^3}} \quad (5.31)$$

The difference between this strategy and TIA+CMD is that using global feedback TIA+CMD the maximum flatness GBW bound of (5.31) can not be realized, while for the case of TIA+CMD the GBW bound of (5.25) can be easily approached. Thus the TIA+CMD design is more bandwidth effective than the global feedback TIA+CMD design.

5.2.4 Delayed Feedback

It is possible to apply the delayed feedback concept to the aforementioned design strategies to improve frequency response. For the case of TIA+CMD, the cutoff frequency can be improved introducing a delay element on the feedback loop of the TIA stage. However, this must be done with some care, since the two dominant poles of the TIA stage interact with the other two poles arising from input circuitry and CMD. The case of the global feedback TIA+CMD design renders

itself more suitable to delayed feedback. For this case, all the four poles are under the feedback action. Furthermore, both CMD and TIA may include RHP zeros (if not, these can be incorporated), leading to easier implementations of the necessary delay. Nevertheless, stability margins may pose severe restrictions on the overall design, since both gain and bandwidth are for this case constrained by a 4th order quasi-polynomial. As discussed on chapter 3, the gain and phase margins showed rapid degradation as the order of the quasi-polynomial increases (see figures 3.14, 3.15 and 3.16).

5.3 Current Conveyors

There are several choices to implement a CMD. Many current amplifier circuits are able to synthesize the CMD's requirements. Simple current amplifiers based on current mirrors (bidirectional or not) present some drawbacks, namely: i) it is difficult to dynamically control the input impedance; ii) it is difficult to set the input bias point; iii) current-mirrors have poor noise performance [290]. One obvious alternative is to use current conveyors. Current conveyors are true current-mode universal circuit elements. As voltage-mode operational amplifiers (OA), current conveyors allow the realization of any kind of transfer function, without recurring to feedback topologies³². They include on the same element both a voltage follower and a current follower (more generally, a current amplifier). In this sense they can be used as current amplifiers and at the same time they are able to overcome the aforementioned drawbacks of current mirrors.

The first CC was originally reported by Sedra and Smith in 1968 [215]. This first CC realization became the first generation CC (or CC-I for short). Since CC-I pose some limitations on the kind of circuits that could be realized (they were mainly devoted to the synthesis of negative impedance converters (NIC) [215]), Sedra and Smith proposed in 1970 another type of CC [216]. The second generation CC (CC-II) allowed a wider range of applications to be designed, and presented the first viable alternative to OAs. Later on, in 1995, a third revision of the CC, the CC-III, was proposed by Fabre [227]. This does not by far complete the range of different CCs types. Several improvements concerning differential signal processing capabilities, both in voltage and current domains have been reported [259, 273, 276, 278, 281]. Nevertheless, these proposals have always evolved from the three basic types of CCs.

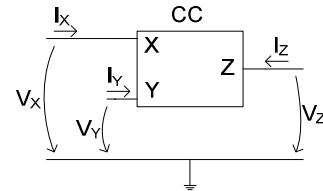


Figure 5.8 - Generic three-port CC representation.

³² Recall that the two basic voltage-mode configurations using OAs, inverting and non-inverting, require feedback. Other kinds of transfer functions are usually derived from these two configurations, thus requiring also similar feedback schemes.

5.3.1 Current Conveyor Types

Figure 5.8 depicts the generic CC three-port representation. As an OA, the CC has in its most general form three terminals, two inputs, X and Y, and one output, Z. CC classification is based on the specific interaction between these three terminals. The output Z serves always as a current output feeding one replica of the input current measured at port X. However, the interactions between the X and Y inputs are different for all the three basic CC types. Considering that each of the three ports in the CC present a driving point impedance z_x , z_y and z_z respectively for the ports X, Y and Z, and that there are no coupling terms between each pair of terminals except those implicitly defined for the correct operation of the CC, the following equation describes all the three types of CC,

$$\begin{bmatrix} I_Y \\ V_X \\ I_Z \end{bmatrix} = \begin{bmatrix} y_y & b_i & 0 \\ a_v & z_x & 0 \\ 0 & \pm a_i & y_z \end{bmatrix} \begin{bmatrix} V_Y \\ I_X \\ V_Z \end{bmatrix} \quad (5.32)$$

In equation (5.32),

- a_i represents the current gain between the X input and the output Z, its value can be either positive or negative or even larger than unity;
- b_i represents the current gain between the inputs X and Y. According to the conveyor type b_i can take different values – $b_i=1$ in a CC-I, $b_i=0$ in a CC-II and $b_i=-1$ in a CC-III;
- a_v represents the voltage gain between the inputs Y and X, which ideally has a unitary value.
- y_y and y_z are the driving point admittances at ports Y and Z, obtained as the reciprocals of the previously defined z_y and z_z , respectively.

The above description reveals that in an ideal CC-I there are precise voltage and current dependencies between the input ports, such that the current flowing in port Y is an exact replica of the current sensed at port X; and also that the voltage developing at port X follows the voltage sensed at port Y [215, 216, 223]. This particular voltage-current dependency is useful for the synthesis of NICs. NICs find several applications on the synthesis of inductorless active filters, where they are often applied as inductance simulators [215, 216]. Other usages of the CC-I resort to current buffering in high frequency applications, such as current probing meters [223, 227].

CC-IIIs have similar properties to CC-Is, except for the current dependency between the inputs X and Y. In an ideal CC-III, the current flowing in the Y port is an inverted replica of the current sensed at port X. The voltage relation remains the same [227]. This particular property of the current inversion is especially important for the synthesis of current measurement devices [227].

Using a CC-I the output current follows in the opposite direction of the original current under measurement. In order to minimize the impact of the measuring device on the circuit, the CC-I has to be designed with current inverting capabilities in its Z output port. The Y input can be used to set the adequate bias voltage and the X input is used to sense the current flowing in the circuit. The major drawback of this scheme is that the output voltage at port Z depends entirely on the impedance of the measuring circuit and may differ from the bias voltage imposed by the Y port. This drawback is effectively removed if a CC-III is used instead. With a CC-III, the current measuring is taken at X and Y ports, where the current inversion is an implicit characteristic.

CC-IIs are more general than CC-Is and CC-IIIs. They can be also applied in the aforementioned examples, NICs and current measurement devices, and too many more analog processing circuits. This general character is apparent from the equivalent circuit description depicted on Figure 5.9. The basic difference between CC-II and the other CCs is that the Y port acts as an ideal high input impedance port, having a negligible input current.

This leaves only the current transfer capability between port X and the output port Z, and voltage dependency of the input X on the voltage at port Y. Ideally, both these current (a_i) and voltage (a_v) gains are unity and

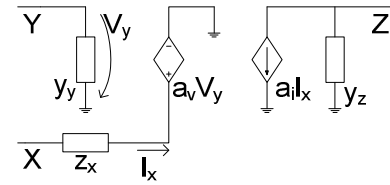


Figure 5.9 - CC-II small signal equivalent circuit.

the final result is a single element able to cope with all the current-mode design strategies discussed in Figure 5.1 and Figure 5.2. This is also the reason for the rapid and ever growing search for new CC-II topologies, since their early invention in 1970 [216]. The equivalent circuit of Figure 5.9 represents three terminal impedances (admittances) z_x , y_y and y_z , representing the driving point impedances (admittances) at the three ports X, Y and Z, respectively. Ideally these impedance and admittances are all zero. However, in real applications their ideal values can only be approached. It has been argued for long time which is the best representation for these impedances [222, 230, 232, 239, 259, 263, 276, 278, 281]. It is accepted that both y_y and y_z can be modeled precisely as a first order parallel association of a resistance and a capacitance (r_z and c_z for y_z , and r_y and c_y for y_y). It is also usually insufficient to consider a simple resistance-capacitance association for z_x as in the previous cases. In general, the high frequency roll-off of z_x approaches the roll-off of a second order driving point impedance. It is also frequent to observe frequency peaking phenomena in z_x , which suggest the presence of complex poles.

The CC-II description is also similar to the description of the generic transistor; in fact, even the terminal labels are the same. This should not come as a surprise, because a single transistor is the simplest implementation of the CC-II. It is sufficient to recall the common Z configuration, having nearly unitary voltage gain from the input (terminal Y) to the voltage output (at terminal X), and

almost unitary current transfer from the input (at terminal X) to the output (at terminal Z). CC-IIs can improve some aspects of this simple transistor description, namely: i) the accuracy of the voltage and current gains is strongly dependent on the load impedances connected to the transistor's terminals; ii) using one transistor there is always a DC voltage drop from terminal Y to terminal X (that is, when the transistor is operated in its linear region – saturation for FETs and forward active mode for BJTs); iii) the terminal impedances may not be satisfactorily close to the ideal situation.

A simple improvement of this simple CC-II is the well known super-transistor [259]: a transistor and a large gain OA connect in a loop comprising the transistor's Y and X terminals (as depicted on Figure 5.10). This configuration uses a feedback scheme to improve both the terminal impedances seen at ports X and Y and also the voltage gain. Due to the high gain of the OA, the necessary voltage drop V_{YX} of the transistor is effectively removed from the input. There are many other possibilities of employing feedback to ameliorate some aspects of the CC-II [295]. Some of these techniques will be explored in section 5.5.

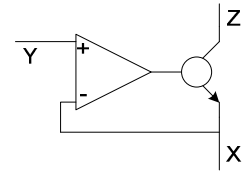


Figure 5.10 - Super-transistor.

5.3.2 Current Conveyor Based CMD Design

The above discussion presented the three types of CCs. It was seen that all of these types can be employed in one manner or another, as current measuring devices. Each of these implementations presents its own characteristics. Figure 5.11 shows the most immediate implementations of CMD using the three types of CC. All these implementations employ a CC with two outputs, one with positive (Z_+) and the other with negative (Z_-) current gain. To gain some insight on the target applications, all these examples include a photo-detector able to convert optic signals into electric currents. One requirement on the CMD is to have a low impedance current input in order to minimize the effect of the photo-detector's intrinsic capacitance. This feature can be easily implemented using the X input of any kind of CC. The particularities of each implementation are intimately related with the function of the second input of the CC. Since the voltage at input X is a replica of the voltage at input Y for all kinds of CCs, the Y input port can serve as a bias voltage

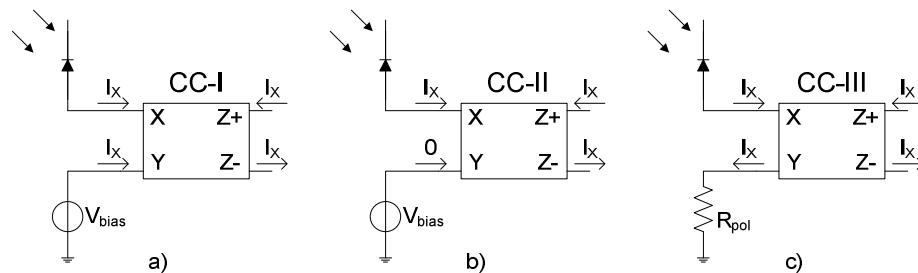


Figure 5.11 - Current measuring using: a) CC-I, b) CC-II, c) CC-III.

input. In this sense, the best possible implementation of the CMD uses a CC-III. The CC-III can act as an ideal current probe, inserted in series with the photo-detector. In its original form, the photo-detector is connected in series with a bias resistor and a reverse voltage source. The major advantage of the CC-III CMD's implementation is exactly the fact that the original bias circuit can be preserved. CC-I or CC-II implementations break this original bias loop and need to incorporate an external voltage source to establish the bias point.

However, the usage of a CC-II can present some advantages to CC-I and CC-III implementations. Since for both CC-I and CC-III, the current flowing in (out, respectively) of terminal Y is a replica of the measured input current, this means that power consumption within the CC is directly proportional to the measured current, which may not be a desired or even tolerated property. In a CC-II this power proportionality is also present in the output stage, but not in the Y input branch. In conclusion, CC-I and CC-III may pose severe power penalties when used as CMD.

The CC-II has further advantages, namely:

- i) They can easily replicate both a CC-I or a CC-III [216], as shown on Figure 5.12. A CC-I is accomplished using one extra positive gain output connected to the Y input. The CC-III has two different implementations. Using a unique CC-II with both positive and negative gain outputs, with the negative output connected to the Y input. If the available CC-II has only positive gain outputs, then it is necessary to use two equal CC-II to implement a CC-III, according to Figure 5.12c;
- ii) They allow superior control of the input impedance present at port X, using feedback control schemes [226, 230, 231, 233, 295];
- iii) They can incorporate differential signal handling capabilities - using two CC-IIs, one for each input current, and combining the outputs accordingly [287, 288, 289], as shown on Figure 5.13; and
- iv) The reported work dealing with CC-II circuit's conception and modeling is more mature than for other CC generations.

These advantages clearly show the superiority

of the CC-II for the design of CMDs. For these reasons the CC-II was selected as the basic element for CMD design on the ambit of this work.

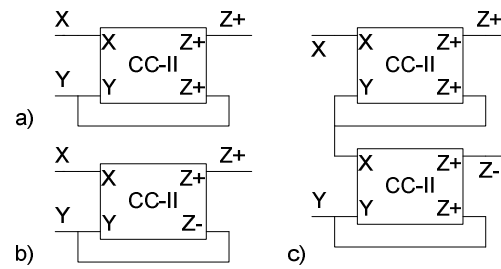


Figure 5.12 - CC-I and CC-III implementations, using the C-II; a) CC-I, b) and c) CC-III.

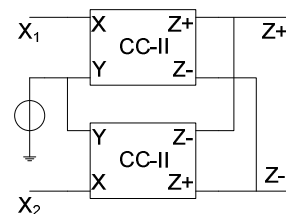


Figure 5.13 – Differential signal handling with CC-II.

5.3.3 Common Topologies

It is rather complicated to categorize CC-IIs. This complication arises due to the large offer of CC-II topologies reflecting different design approaches. Such a categorization would have to include topics like:

- i) Low-voltage adequacy [225, 231, 235, 281];
- ii) Low-power operation [235, 281];
- iii) High-frequency operation [224, 236, 240, 241, 243, 245, 278];
- iv) High-drive capabilities [223, 228, 235, 273, 281];
- v) High-accuracy [223, 224, 228, 235, 273, 278]; and,
- vi) Low-noise behavior [229, 281]; and possibly, many others.

However, all these categories share one particular characteristic in common, the possibility of exhibiting large or small current gain, which suggest a rather simple and more circuit oriented classification. Under this classification, there are two types of CC-II's topologies: small gain and large gain configurations. From the above discussions it is clear that the condition of unitary voltage gain in a CC-II is a required condition for the functionality of the device. However, nothing restricts the magnitude of the current gain in a CC-II. This observation brought an opportunity for the conception of large gain CC-II topologies [278]. Nevertheless, it should be noted that large gain configurations do not present as competitive characteristics over traditional OA as small gain configurations. Obviously, large gain comes at the expense of reduced bandwidth.

This section presents a short overview of some circuit concepts behind these two strategies, small gain and large gain. It is not intended to explore in full detail all the available variations of such circuits, which is obviously outside the scope of this text. However, a better comprehension of such concepts has helped to study and improve the CC-II topologies used in this work.

5.3.3.1 Small Gain Configurations

The simplest CC-II implementations stem from the single transistor interpretation previously discussed. These topologies are often associated to circuits operating in class A (having all the transistors in the linear region). Figure 5.14 depicts these simplest class A topologies. Figure 5.14a and b represent positive and negative current gain versions of the same CC-II. Only one transistor (T_1) is needed to implement the functionalities of the CC-II. Transistor T_1 synthesizes the voltage relation between the inputs Y and X. Its high impedance terminal is the Y input of the conveyor. The X input is connected to the X terminal of T_1 , where the driving point impedance is very small. The output is taken as mirrored replica of the current flowing into de Z terminal of T_1 . Positive gain implementations may exhibit non unitary current gains, due to the presence of a current mirror; the

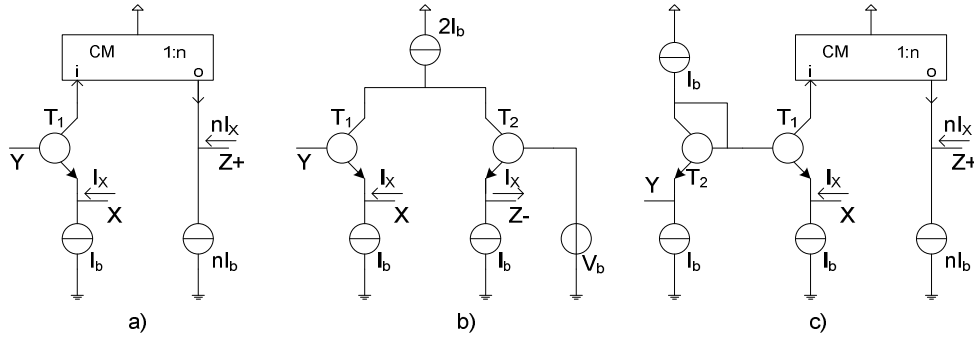


Figure 5.14 - Class A CC-II circuit topologies: a) simple one transistor positive gain CC-II; b) negative gain CC-II; c) improved positive gain CC-II.

current gain can be adequately adjusted with the design ratio of the transistors in the current mirror. However, large current gains imply large asymmetries in the transistors of current mirror, which usually traduces in reduced performance (low accuracy, severe bandwidth penalties, increased equivalent noise [290, 291]). Solutions to this problem resort to: i) improved current mirror topologies to combat accuracy problems, implying larger complexity; ii) cascaded current mirror topologies to alleviate the bandwidth problems [291], but this corresponds to degraded noise performance. In fact, from the noise point of view, it is always better to use direct mirror conversion than cascaded topologies [290]. In this respect, negative gain CC-II present superior performance. As depicted in Figure 5.14b, a negative CC-II does not require current mirrors to derive the output. Instead, a simple level-shifter, T_2 and an adequate voltage reference are used to this end. Unfortunately, negative CC-IIs like the one depicted in Figure 5.14b have a stringent restriction on gain: the magnitude of the current gain is always unitary.

Figure 5.14c presents an improved version of the CC-II of Figure 5.14a. One of the major drawbacks of the simple circuit of Figure 5.14a lies on the fact that there is always a necessary voltage drop between the terminals Y and X to maintain T_1 functioning in its linear regime. This voltage drop can be compensated inserting an auxiliary level-shifter as in Figure 5.14c. Transistor T_2 performs the adequate level-shifting; if the quiescent conditions are similar for both T_1 and T_2 , then, the inputs X and Y have similar DC bias point. Other alternatives to accomplish the same effect resort to the usage of differential pairs, as in Figure 5.15. These configurations are also implementations of the super-transistor concept previously discussed (see Figure 5.10). In fact the differential pair formed by transistors T_1 and T_2 is a simple implementation of the OA. There are many CC-II structures using differential pairs, but they all resort to the same concepts of Figure 5.15. A potential superiority of these configurations is their inherent ability to provide both negative and positive gain; it suffices to sample the output current on the adequate branch of the differential pair. Improvements on these configurations resort to: i) larger gain differential pairs; ii) rail-to-rail capabilities; iii) efficient bias schemes able to cope with low-voltage constraints; and iv) operational

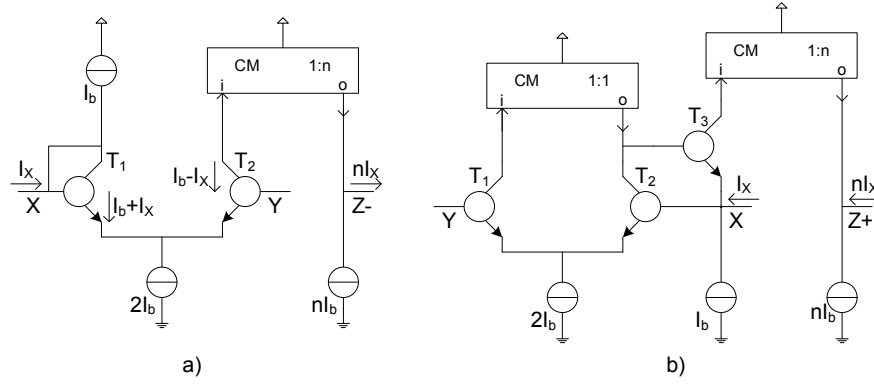


Figure 5.15 - CC-II topologies based on the differential pair: a) negative gain; b) positive gain.

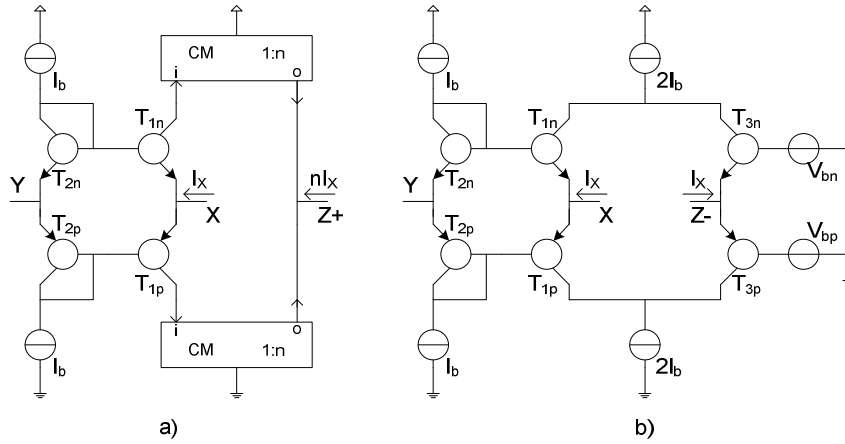


Figure 5.16 - Class AB CC-II structures: a) positive gain; b) negative gain.

transconductance amplifiers (OTA) based configurations. The major advantages of these configurations are the enhanced accuracy of the resulting voltage transfer gain and the reduced X input impedance. Nevertheless, the usage of differential pairs can represent serious bandwidth limitations. High bandwidth applications often resort to the usage of simple class A or class AB CC-II's structures without such improvements.

Class AB structures are derived from their class A counterparts. The most common topologies are derived from the circuits of Figure 5.14a and b, adding the level-shifting improvement. Figure 5.16 depicts two class AB CC-IIs, having positive and negative gain outputs. It is easily seen the resemblance between these circuits and the circuits of Figure 5.14; in fact, isolating each group of transistors of the same type, these circuits are just the parallel combination of one N type class A CC-II with a P type class A CC-II (except for the bias current sources that are no longer necessary for the positive CC-II). Class AB CC-IIs present some advantages over their class A counter parts, namely: i) improved accuracy on both current and voltage gain; ii) larger current input/output dynamic range; iii) efficient bias control; and iv) improved immunity to temperature and process variations. The true essence of these improvements stems from the fact that the four transistors that comprise the input stage of the conveyor (T_{1n} , T_{2n} , T_{1p} and T_{2p}) form what is commonly known as a

translinear loop. Translinear loops were original introduced in bipolar technology by Gilbert in 1975 [276], and later generalized to CMOS by Wiegink in 1991 [265]. Translinear loops have the property of allowing the implementation of exact, temperature and process insensitive mathematical functions. Class AB CC-IIs explore in full extent this property, not to directly implement a mathematical function, but to combine the result of these functions in such a way to provide adequate CC-II's functionalities. These features make class AB CC-IIs based on the translinear principle one of the preferred CC-II's topologies. These reasons lead this work to adopt this type of CC-II. One minor drawback of these structures is that they require complementary processes, providing both N and P type devices with good characteristics. For this reason, most of the reported work around class AB CC-IIs uses CMOS processes, since bipolar processes are usually not complementary.

5.3.3.2 Large Gain Configurations

It is possible to convert almost all of the described small gain CC-II's configurations into large gain configurations. The concept underlying this transformation is briefly depicted in Figure 5.17. Small gain configurations use current mirrors to convey a replica of the input current (entering at

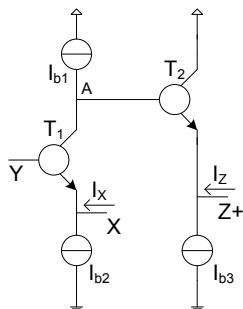


Figure 5.17 – Large gain CC-II: design concept.

port X) to the output terminal Z. Increasing the gain on such configurations is achieved using large gain current mirrors or cascade associations of current mirrors (whenever possible). But even with such arrangements the achieved current gain is still very modest when compared with the voltage gain of OAs. A straight solution to increase even further the current gain is to rely on active devices. Transistors have generally large current gains (MOSFET transistors have in this respect almost infinite

current gain, while BJT exhibit smaller values). Instead of using current mirrors to convey currents between input and output, it is possible to use the magnifying properties of the transistor itself to form the output current. Such arrangement is represented in Figure 5.17. The current I_X sets the value of the voltage at node A (a high impedance node), which in its turn controls the Y terminal (gate on a MOSFET or base in a BJT) of the output transistor T_2 .

These large gain configurations have an inherent drawback: limited bandwidth. Obviously, increasing gain leads to smaller bandwidths as usual. Nevertheless, it is instructive to inspect what is the reason for bandwidth loss in this case. As mentioned above, there is a current to voltage conversion on the high impedance node A. This node sees an equivalent load resistance consisting on the parallel association of the input resistance of T_2 and the output resistance of a current source.

In general this represents a very large resistance, which magnifies the effect of any small capacitance arising at this node. The associated time constant with this node is the principal bandwidth limitation factor on large gain configurations. Large gain topologies contradict the basic principles of current-mode design techniques. As said previously, high impedance nodes are something to avoid in current-mode design.

5.4 Positive Class AB CC-II Design Considerations

Class AB CC-IIs based on translinear loops have demonstrated several interesting properties: enhanced dynamic ranges, suitability for low voltage operation, high temperature immunity, good accuracy, large bandwidth, enhanced power control dynamics and uniform layout implementations. These characteristics make them suitable for analog signal processing applications, especially under low-voltage constraints.

The following discussion will be restricted to CMOS class AB CC-II current conveyors based on the translinear principle. Thus the generic transistor model used until now throughout this text is replaced by the more suitable model of the MOS transistor. Three aspects of the conveyor modeling are discussed: current and voltage limitations in the static regime, and frequency response.

5.4.1 Current Dynamic Range Limitations

The most important characteristic on current conveyors is their ability to convey currents between input (X) and output (Z) terminals. Thus it is of paramount importance to have a clear picture of the current conveying mechanisms, which for these type of current conveyors are mainly governed by the input circuitry. Figure 5.18 represents the input stage of a CMOS class AB CC-II based on one translinear loop. The output stage is simply formed by the adequate combination of the tail currents I_N and I_P , as shown on Figure 5.16. For now the attention is focused on the input stage. Transistors M_1 , M_2 , M_3 and M_4 form a translinear loop [259, 265, 276]. Since transistors can

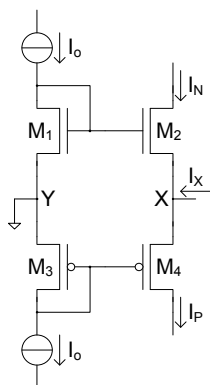


Figure 5.18 – Class AB CC-II input stage.

be modeled as voltage controlled current sources, combining transistors in loop arrangements such that all the control voltages become constrained by Kirchoff's voltage law, allows to convert voltages equations into useful current relations (using the transistor's characteristics). Furthermore, these current relations provide temperature insensitive means of achieving mathematical relations of profound impact on circuit design. Originally, the translinear principle was applied to bipolar transistors only. The concept was later generalized by Wiegink to MOS transistors [265].

Returning to the circuit of Figure 5.18, the translinear loop is formed by the gate to source voltage drops of transistors M_1 to M_4 . Assuming that all the transistors operate under saturation regime, and that the drain current is given by,

$$I_D = k(V_{GS} - V_T)^2 \quad (5.33)$$

where, k represents the transistor's transconductance parameter (given as $0.5\mu C_{ox}(W/L)$, where μ is the mobility of the transistor – electrons in NMOS transistors and holes in PMOS transistors; C_{ox} is the capacitance of the gate oxide and (W/L) is the aspect ratio of the transistor), V_{GS} is the gate to source voltage and V_T the threshold voltage. Assuming also that the transconductance parameters are k_n and k_p for NMOS and PMOS transistors respectively, applying the translinear principle leads to,

$$\begin{aligned} V_{GS_1} + |V_{GS_3}| &= V_{GS_2} + |V_{GS_4}| \\ V_{T_{n1}} + \sqrt{\frac{I_o}{k_n}} + |V_{T_{p3}}| + \sqrt{\frac{I_o}{k_p}} &= V_{T_{n2}} + \sqrt{\frac{I_N}{k_n}} + |V_{T_{p4}}| + \sqrt{\frac{I_P}{k_n}} \end{aligned} \quad (5.34)$$

Assuming that the threshold voltages have nearly the same magnitude for both transistor types and introducing a new design variable r , given by k_n/k_p , equation (5.34) simplifies into,

$$(1 + \sqrt{r})\sqrt{I_o} = \sqrt{I_N} + \sqrt{rI_P} \quad (5.35)$$

The temperature dependent terms are mainly associated with the threshold voltage and the mobility of the transistors. The threshold voltage are cancelled directly on (5.34). Compensation of the temperature effects associated with the mobility factor require that $r=1$ to cancel out. Combining (5.35) with the equation for the currents at node X, $I_N + I_X = I_P$, allows to find general solutions for I_N and I_P . It is possible to express these solutions in a mathematical form. However, these forms are not simple to read. One particular case of interest is when r is unity. For this case, I_N and I_P are given by,

$$I_{N,P} = I_o \left(1 \mp \frac{I_X}{4I_o} \right)^2 \quad (5.36)$$

Figure 5.19 shows normalized versions of the currents I_N and I_P taking the normalized input current, I_X as independent variable. It is apparent the symmetry between I_N and I_P . Also apparent from Figure 5.19 and equation (5.36) is the input dynamic range for I_X which covers all the current range from $-4I_o$ to $4I_o$. The maximum tail current is also equal to $4I_o$. These observations suggests that the biasing current sources I_o furnish an efficient mean of controlling both power consumption and input dynamic range.

This scenario changes slightly when r takes other values different than unity. For this case, I_N

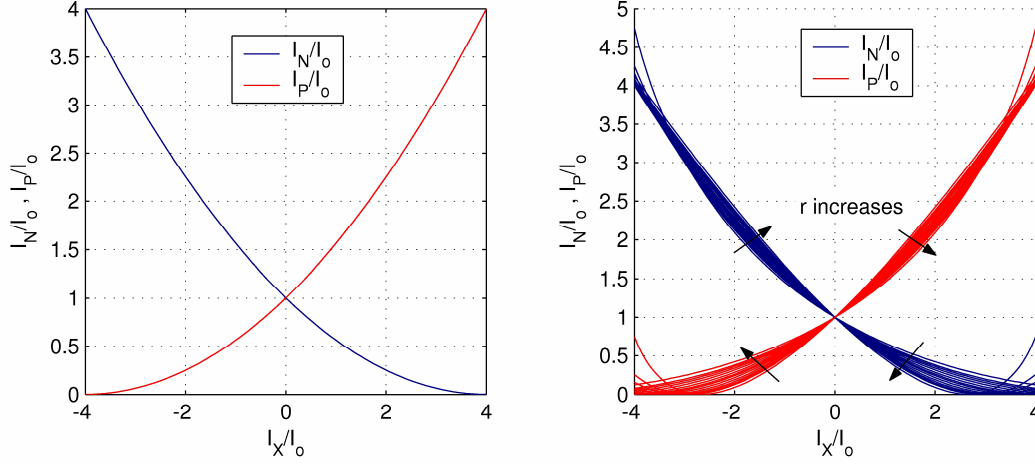


Figure 5.19 – Normalized currents I_N and I_P (matched transconductance case, $r=1$, on the left; changing r for several values between 0.4 and 2.5 , on the right).

and I_P lose symmetry. Figure 5.19 shows the effect on I_N and I_P when r takes several values between 0.4 and 2.5 . It is readily seen that the maximum current can surpass $4I_o$ and that the input dynamic range shifts from right to left (the arrows on Figure 5.19 show the turning of the current characteristics, when r increases). Nevertheless, for values between 0.4 and 2.5 , these changes are not as dramatic as it would be expected. Thus it is possible to use different aspect ratios for the NMOS and PMOS transistors, in order to promote some other desired characteristic (for instance, enhanced frequency response), rather than the usual aspect ratio needed to equalize the transconductance parameters ($r=1$). However, this should be done with special care in order to cope with input dynamic range requirements.

To find the general dynamic range limitations valid for all r values, it is necessary to observe carefully Figure 5.19. For $r=1$, I_N and I_P are two parabolas with vertical symmetry, having a minimum that coincides with the zero current value. This minimum specifies the input dynamic range limitations ($\pm 4I_o$). For the general case the same observations are also true; the difference is that now the two I_N and I_P parabolas have diagonal symmetry. Nevertheless, their minimum still coincides with the zero current value. Solving simultaneously (5.35) and $I_N + I_X = I_P$ in order to I_P and inserting the condition $I_P=0$ on the final result, returns the lower bound of the input current I_X , given by I_1 in (5.37).

$$I_1 = -I_o (1 + \sqrt{r})^2 \quad (5.37)$$

Proceeding in a similar manner in order to I_N , returns the upper bound I_2 , given in (5.38).

$$I_2 = I_o \frac{(1 + \sqrt{r})^2}{r} \quad (5.38)$$

In general the input dynamic range is limited on the two extremes by I_1 (on I_P side) and I_2 (on I_N

side). Figure 5.20 represents the normalized limits on I_X , I_I and I_2 . As can be seen, small variations of r near unity produce almost symmetric shifts on the dynamic range. However, when r is larger or smaller than unity, the dynamic range shift becomes very asymmetric.

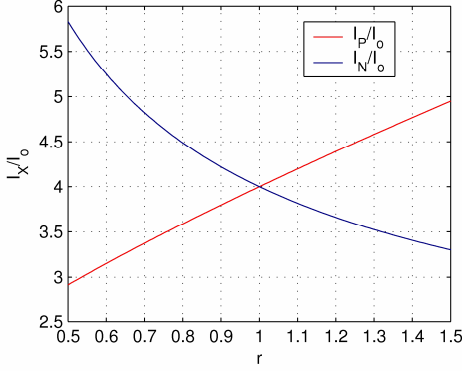


Figure 5.20 – Dynamic range dependence on r .

The output current dynamic range follows directly from the input limitations. In a positive gain CC-II like the one represented on Figure 5.16a, the output current is an amplified version of the I_X . The amplification factor is adjusted by the gain of the current mirrors used to produce the output replicas of I_N and I_P ; for the case of Figure 5.16a this results in $I_Z = nI_X$, where n represents the current gain. This also means that the output current range can reach n times the input range, that is, from nI_I to nI_2 . Increasing the gain leads to large output dynamic ranges, but also implies larger power consumption.

5.4.2 Voltage Dynamic Range Limitations

The previous section discussed the current limitations found in class AB CC-II based on the translinear loop principle. However, these are not the unique limitations for this kind of CCs. The previous results are valid under the assumption that all transistors operate on the saturation regime, allowing the CC-II to operate in its linear range. Whenever these assumptions fail, the transistors leave saturation and operate in the triode region or even enter the cut off mode. For these situations the CC-II becomes highly nonlinear, and the terminal voltages approach their saturated values, near the bias voltages rails.

The voltage limitations are found when the transistors reach the boundary between saturation and triode regions. This boundary is roughly defined when the drain to source voltage drop, V_{DS} , reaches the limit value $V_{GS} - V_T$, that is,

$$V_{DS} = V_{GS} - V_T = \sqrt{I_D/k} \quad (5.39)$$

Equation (5.39) allows to investigate the maximum voltage ratings on CC-II. Figure 5.21 depicts all the possible transistors associations in the branches of a translinear CC-II, using simple current mirrors (CM), and one transistor current sources (CS). Starting with the Y branch (Figure 5.21a), the maximum voltage span possible in the Y node has two limits: the upper limit imposed by the bias voltage V_{DD} , the V_{DS} of transistor M_7 on the limit of saturation and the V_{GS} of M_1 ; and the lower limit imposed by the bias voltage V_{SS} , the V_{DS} of transistor M_8 on the limit of saturation and the V_{GS} of M_3 . Using (5.33) and (5.39), this limits become,

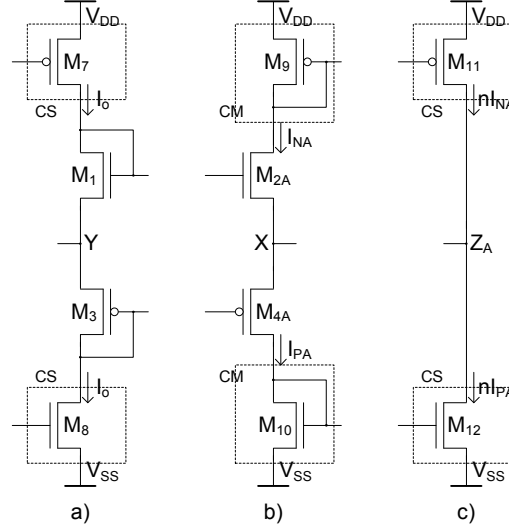


Figure 5.21 – X, Y and Z branches of a CC-II: a) Y branch with simple current source (CS); b) X branch with mirroring transistors (CM); and c) Z branch with CS.

$$\begin{aligned} V_{Y_{\max}} &= V_{DD} - V_{Tn} - \left(\sqrt{1/k_1} + \sqrt{1/k_7} \right) \sqrt{I_o} \\ V_{Y_{\min}} &= V_{SS} + |V_{Tp}| + \left(\sqrt{1/k_2} + \sqrt{1/k_8} \right) \sqrt{I_o} \end{aligned} \quad (5.40)$$

The X input has slightly different voltage limitations, since now there is another current contribution. The drain currents of the transistors M_{2A} and M_{4A} in Figure 5.21b are now dependent on the input current I_X . Applying the same reasoning as for the case of the Y branch, the X branch limitations are for the first case,

$$\begin{aligned} V_{X_{\max}} &= V_{DD} - |V_{Tp}| - \left(\sqrt{1/k_{2A}} + \sqrt{1/k_9} \right) \sqrt{I_{NA}} \\ V_{X_{\min}} &= V_{SS} + V_{Tn} + \left(\sqrt{1/k_{4A}} + \sqrt{1/k_{10}} \right) \sqrt{I_{PA}} \end{aligned} \quad (5.41)$$

Finally, the Z branch (see Figure 5.21) voltage limitations are given by,

$$\begin{aligned} V_{Z_{\max}} &= V_{DD} - |V_{Tp}| - \sqrt{nI_{NA}/k_{11}} \\ V_{Z_{\min}} &= V_{SS} + V_{Tn} + \sqrt{nI_{PA}/k_{12}} \end{aligned} \quad (5.42)$$

From equations (5.40) and (5.41) it is readily seen that the X and Y nodes are the ones with stringent voltage constraints. For these cases, the voltage limits can reach closely to one threshold voltage under the bias rails. There is also a small influence posed by the branch current; in particular, for the case of the X branch (and also the Z branch) these current dependence is related to the current entering the X input (Z output), which may represent an additional detrimental factor.

5.4.3 Frequency Response Optimization

The frequency response of an class AB CC-II can be described using two class A half circuits. The concept behind class AB circuit conception uses two class A complementary circuits associated

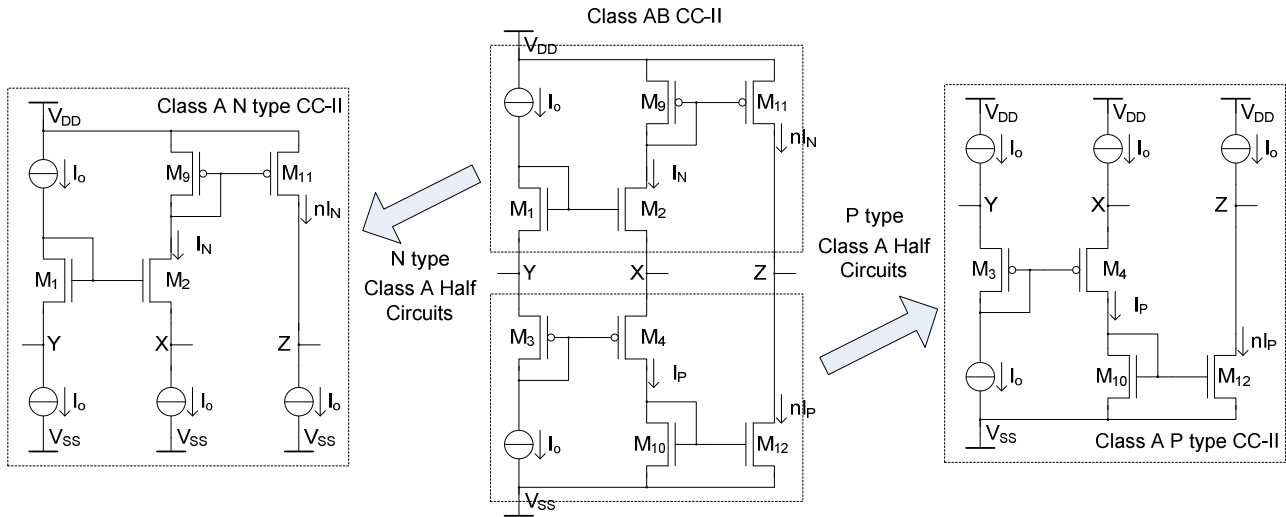


Figure 5.22 – Class AB Positive CC-II half circuits.

in a parallel fashion as shown in Figure 5.22. In order to study the frequency behavior of the complete class AB CC-II, it is possible to consider independently each of this class A half circuits. This simplification traces the path to optimized frequency response of the overall circuit. If the two class A half circuits have precisely (or at least nearly) the same frequency behavior, differing only on magnitude terms, the complete class AB circuit has also the same frequency behavior as its class A half circuits. The justification of this fact can be formalized using an admittance matrix representation. Using standard nodal analysis, each class A half circuit has an associated admittance matrix of the form,

$$\begin{aligned} [I] &= [y_{ij}^N(s)]_N [V] \\ [I] &= [y_{ij}^P(s)]_P [V] \end{aligned} \quad (5.43)$$

where the matrixes $[I]$ and $[V]$ are column vectors representing the node input currents and node voltages respectively. The subscripts N and P (superscripts in $y_{ij}(s)$) refer to the N and P class A half circuits and $y_{ij}(s)$ represents the elements of the admittance matrix. Now, assuming that under quiescent conditions each element $y_{ij}^N(s)$ has the same frequency behavior as $y_{ij}^P(s)$, that is,

$$\frac{y_{ij}^N(s)}{y_{ij}^N(0)} = \frac{y_{ij}^P(s)}{y_{ij}^P(0)} = h_{ij}(s) \quad (5.44)$$

the overall class AB admittance matrix is just the sum of the two admittance matrixes in (5.43), since the complete circuit is assumed to be obtained as a parallel association of the two N and P class A half circuits, that is

$$[I] = [(y_{ij}^N(0) + y_{ij}^P(0))h_{ij}(s)][V] \quad (5.45)$$

There are some intricacies that have to be explained in order to achieve the form of (5.45). First of all, equation (5.45) implies a frequency matching procedure, expressed mathematically in (5.44).

It is not simple to achieve this idealized frequency matching. Two factors contribute to this difficulty. First, it would be required to have PMOS and NMOS transistors with matched transition frequencies and equalized transconductance to output conductance ratio. These conditions are difficult to fulfill simultaneously since both factors depend on the device transconductance. Second, when the circuit is operated under normal input signal conditions, the transistors on each branch experiment deviations from the idealized quiescent conditions. According to Figure 5.19, this implies that some transistors are operated with larger currents than others. Since the transition frequency is strongly dependent on the bias current, this also means that some transistors become faster than others. Nevertheless, this behavior can be considered approximately symmetrical around the quiescent condition. So the two half circuits will tend to compensate each other; when the N half circuit becomes slower the P half circuit performs faster and vice-versa.

5.4.4 Class A Half Circuit Normalized Frequency Response

It is difficult to formally express the complete frequency response of a CC-II. The reason for this difficulty is readily apparent. Circuits containing more than two or three transistors have extremely complex dynamics; since each transistor contributes with two different intrinsic capacitances, the complete frequency response contains polynomial factors of increasing order. Furthermore, since the time constants associated with each frequency power are directly related to the intrinsic capacitances and to some special circuit driving point impedances [87, 89, 294], the circuit transfer functions become extremely complex to read. Nevertheless, it is important to have a clear theoretical perspective of these transfer functions prior to design and simulation.

It is possible to simplify the analysis imposing some early design constraints known to promote some desired behaviors. For instance, for the case of class AB CC-II design the first simplifying assumption is to divide the circuit into two class A half circuits as seen above. Another powerful tool in this sense is to use normalized frequency response transfer functions to represent the circuit's nodal admittance matrix. A possible frequency normalization is to use the transition frequency of the transistors as normalizing constant.

Considering a positive class AB CC-II, the associated class A N half circuit simplified small signal equivalent circuit is depicted in Figure 5.23. It is readily apparent that the tail current sources connected to the Y, X and Z terminals were neglected (see Figure 5.22). Since, the complete class AB circuit has no tail current sources connected to these terminals, their presence on the class A half circuits can be considered ideal (thus meaning, ideal open-circuits for small signal analysis). The other current source (CS, represented by r_s and c_s) is part of the class AB circuit and must be considered for completion purposes. It will be assumed that the NMOS transistors M_1 and M_2 , and

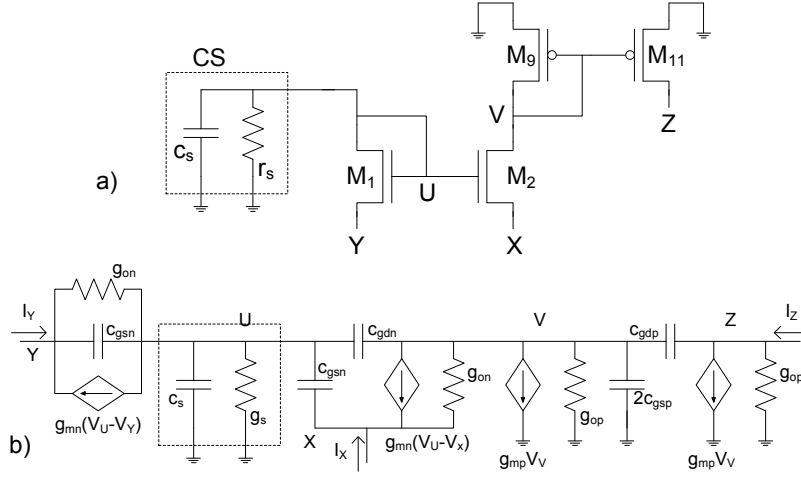


Figure 5.23 – Positive class A CC-II N half small signal equivalent circuit: a) simplified version; b) complete version.

the PMOS transistors M_9 and M_{11} are all operated under saturation conditions at the same bias current I_o and have pair wise identical small signal equivalent models; same transconductance g_m , same output conductance g_o , and same intrinsic capacitances c_{gs} and c_{gd} , defined by,

$$\begin{aligned}
 g_m &= \frac{\partial I_D}{\partial V_{GS}} = \sqrt{2\mu c_{ox} \left(\frac{W}{L} \right) I_o} \\
 g_o &= \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{1}{\lambda I_o} \\
 c_{gs} &= c_{ox} W (L + L_D) \\
 c_{gd} &= c_{ox} W_n L_D
 \end{aligned} \tag{5.46}$$

where μ , λ , c_{ox} and L_D are parameters representing the electron mobility, the channel length modulation, the oxide capacitance and the gate overlap over the drain and source regions, respectively. W and L are the width and length of the transistor. In order to differentiate the device type, a subscript n or p will be appended to all the above parameters, for N type (M_1 and M_2) and for P type transistors (M_9 and M_{11}), respectively. It will be further assumed that both L and L_D have similar values for NMOS and PMOS transistors. L_D is a model parameter used to account the etching defects during the integrated circuits fabrication, its basic effect on transistor operation is to reduce the effective gate length (L). Applying the above considerations together with standard nodal circuit analysis to the circuit of Figure 5.23b results in the following matrix equation,

$$\begin{bmatrix} I_Y \\ I_X \\ I_Z \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} y_1 & 0 & 0 & -y_1 & 0 \\ 0 & y_1 & 0 & -y_2 & -g_{on} \\ 0 & 0 & y_6 & 0 & -y_5 \\ -y_1 & -sc_{gs_n} & 0 & y_u & -sc_{gd_n} \\ 0 & -y_4 & -sc_{gd_p} & y_3 & y_v \end{bmatrix} \begin{bmatrix} V_Y \\ V_X \\ V_Z \\ V_U \\ V_V \end{bmatrix} \tag{5.47}$$

Almost all of the elements of the admittance matrix in equation (5.47) are frequency dependent.

Their complete definition follows in (5.48).

$$\begin{aligned}
 y_1(s) &= g_{m_n} + g_{o_n} + sc_{gs_n} \\
 y_2(s) &= g_{m_n} + sc_{gs_n} \\
 y_3(s) &= g_{m_n} - sc_{gd_n} \\
 y_4(s) &= g_{m_n} + g_{o_n} \\
 y_5(s) &= g_{m_p} - sc_{gd_p} \\
 y_6(s) &= g_{o_p} + sc_{gd_p} \\
 y_u(s) &= y_1(s) + s(c_{gs_n} + c_{gd_n}) + y_{CS}(s) \\
 y_v(s) &= g_{m_p} + g_{o_p} + g_{o_n} + s(2c_{gs_p} + c_{gd_p} + c_{gd_n})
 \end{aligned} \tag{5.48}$$

The element y_{CS} represent the output admittance of the current source CS.

The first step towards normalization is to identify common elements. Looking to (5.47) and (5.48) it seems adequate to put the terms $g_{m_n} + g_{o_n}$ into evidence. Furthermore, since the transition frequency is given for the NMOS transistors by the reciprocal of $\tau_n = (c_{gs_n} + c_{gd_n})/g_{m_n}$, the intrinsic capacitances c_{gs_n} and c_{gd_n} can be expressed as adequate proportions of τ_n . The same observation is also valid for c_{gs_p} , c_{gd_p} and τ_p . Applying these transformations to (5.47) and (5.48) result in the following normalized admittance matrix,

$$[Y]_N = Y_0^N \left[\begin{array}{c|c} H_p & A \\ \hline B & H_{uv} \end{array} \right]_N = Y_0^N \left[\begin{array}{ccc|cc} h_1 & 0 & 0 & -h_1 & 0 \\ 0 & h_1 & 0 & -h_2 & -h_7 \\ 0 & 0 & h_6 & 0 & h_5 \\ \hline -h_1 & -h_8 & 0 & h_u & -h_9 \\ 0 & -1 & -h_{10} & h_3 & h_v \end{array} \right]_N \tag{5.49}$$

The normalized transfer functions are defined according to Table 5.1. The normalizing factors consist of conductance ratios (γ , η_n , and η_p), capacitance ratios (η_c), one transition frequency ratio (ρ) and some simplifying constants (c , b_n and b_p), which are defined as: $\eta_c = c_{gs_n}/c_{gd_n} = c_{gs_p}/c_{gd_p}$, $\eta_n = g_{m_n}/g_{o_n}$, $\eta_p = g_{m_p}/g_{o_p}$, $\gamma = g_{m_p}/g_{m_n}$, $\rho = \omega_n/\omega_p = \tau_p/\tau_n$, $b_n = \eta_n/(1 + \eta_n)$, $b_p = \eta_p/(1 + \eta_p)$, $c = 1/(1 + \eta_c)$. The normalized frequency variable is taken as $u = s\tau_n$, using the NMOS transition frequency as normalizing factor. The same frequency normalization is also used PMOS devices, using ρ as conversion factor, to convert τ_p into τ_n . Y_0 represents the normalizing impedance given as $g_m + g_o$, for both N and P half circuits. The same normalizing functions are easily transformed to describe the P half circuit frequency response. For that purpose, it suffices to replace every NMOS parameter by its correspondent PMOS and also to replace the factors γ , ρ and the normalized frequency variable u by, $1/\gamma$, $1/\rho$ and ρu , respectively. Table 5.1 exhibits implicitly these transformations, into two columns; one describing the N type class A half circuit, and other describing the P type.

Table 5.1 – Normalized transfer functions.

	<i>Half N</i>	<i>Half P</i>
Y_0	g_{m_n}/b_n	$g_{m_n}\gamma/b_p$
$h_1(u)$	$1 + \eta_c b_n c u$	$1 + \eta_c \rho b_p c u$
$h_2(u)$	$b_n (1 + \eta_c c u)$	$b_p (1 + \eta_c \rho c u)$
$h_3(u)$	$b_n (1 - c u)$	$b_p (1 - \rho c u)$
$h_4(u)$	1	1
$h_5(u)$	$b_n \gamma (1 - \rho c u)$	$b_n (1 - c u) / \gamma$
$h_6(u)$	$\frac{b_n \gamma}{\eta_p} (1 + \eta_p \rho c u)$	$\frac{b_p}{\gamma \eta_n} (1 + \eta_n c u)$
$h_7(u)$	b_n / η_n	b_p / η_p
$h_8(u)$	$\eta_c b_n c u$	$\eta_c \rho b_p c u$
$h_9(u)$	$b_n c u$	$\rho b_p c u$
$h_{10}(u)$	$\gamma \rho b_n c u$	$b_p c u / \gamma$
$h_u(u)$	$1 + b_n (1 + \eta_c c) u$	$1 + \rho b_p (1 + \eta_c c) u$
$h_v(u)$	$b_n \gamma k_v \left(1 + \frac{\rho}{k_v} \left(1 + c \left(\eta_c + \frac{1}{\gamma \rho} \right) \right) u \right)$	$\frac{b_p k_v}{\gamma} \left(1 + \frac{1}{k_v} (1 + c (\eta_c + \gamma \rho)) u \right)$
k_v	$1 + \frac{1}{\eta_p} + \frac{1}{\gamma \eta_n}$	$1 + \frac{1}{\eta_n} + \frac{\gamma}{\eta_p}$

The admittance elements due to the current source, y_{CS} in (5.49), have negligible impact on the frequency response if the correspondent current sources is designed in order to meet $|y_{CS}(ju)| \ll |Y_0|$. The normalized transfer functions were simplified assuming this condition. Nevertheless, it is not difficult to take y_{CS} into account (it suffices to add its normalized transfer function to $h_u(u)$).

Referring to the N type to P type transformations previously described, three facts deserve attention. First the factor η_c defined as the ratio of intrinsic capacitances on a NMOS transistor is not replaced by a similar description for PMOS transistors. Using the definitions in (5.46) η_c becomes,

$$\eta_c = \frac{c_{gs_n}}{c_{gd_n}} = \frac{2}{3} + \frac{L}{L_D} \quad (5.50)$$

Unless PMOS transistors are designed with different gate lengths (L) than NMOS, η_c must be equal for both types of transistors, since it was assumed that L_D as the same values for NMOS and PMOS transistors³³. Second, there is not much freedom to choose adequate values for η_c . According

³³ This is a reasonable approximation due to the similar nature of the etching defects during the producing phase of

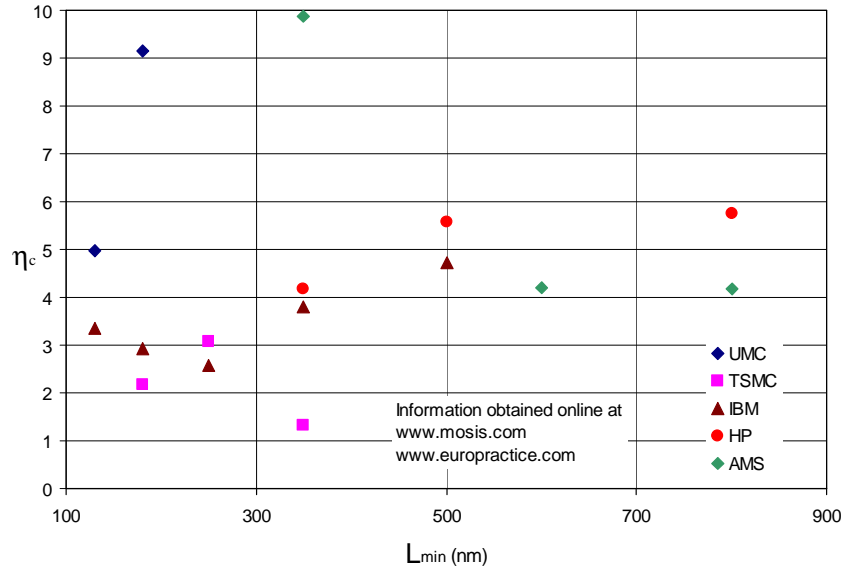


Figure 5.24 - η_c at different technological nodes and different foundries.

to (5.50) η_c depends linearly on L , which may suggest to use different values of L to shape adequately η_c . However, changing L has a dramatic impact on the frequency behavior, since the transition frequency depends inversely on the $L^{3/2}$. In this respect it is reasonable to assume that all the transistors are designed with minimum gate lengths (or at least near the minimum). Figure 5.24 depicts values of η_c for different L_{min} and different CMOS foundries, justifying the assumption that η_c is a technological dependent parameter. Finally, the values of η_n and η_p are naturally different for NMOS and PMOS transistors.

It was suggested in section 5.4.3 that in order to match the frequency response of both class A half circuits, the NMOS and PMOS transistors should have similar transition frequencies and similar output conductances. In fact, according to Table 5.1, matched frequency response requires more than these two conditions. Comparing the normalized transfer functions of both half circuits, reveals that it is necessary to fulfill three conditions in order to have matched frequency response. These conditions are: $\gamma=1$, $\rho=1$ and $\eta_n=\eta_p$. Matching the transition frequency for PMOS and NMOS transistors ($\gamma=1$), imply different gate widths, and consequently different g_m values ($\rho \neq 1$)³⁴. This observation implies that $\gamma=1$ and $\rho=1$ are not mutually realizable and consequently, it is not possible to have matched frequency response. However, it is possible to approximate this condition, but this does not correspond to an optimized frequency response circuit, exhibiting the largest

PMOS and NMOS transistors. However, since L_D is also a model parameter subjected to fitting approximations, it is not surprising to find values of L_D slightly different for both types of transistors.

³⁴ Recall that the necessary condition to equalized the transconductance of PMOS and NMOS devices is $W_N/W_P = \mu_p/\mu_n$.

possible bandwidth. Nevertheless, the bias current and also the gate length of the transistors can be adequately adjusted in order to provide the desired frequency response. It is also noticeable that regardless to the different values of η_n and η_p , the correspondent b_n and b_p can be made approximately the same, using small output conductance devices.

5.4.5 Complete Frequency Response

The complete CC-II's frequency response is defined by the sum of the admittance matrixes of the two class A half circuits. However, to accomplish this summation it is required to express adequately each admittance matrix. This conversion step is a necessary one, since the matrix equation in (5.49) includes information about the half circuit internal nodes. Considering the matrix partition expressed in (5.49) this conversion is accomplished using,

$$[Y_T]_N = \frac{g_{m_n}}{b_n} \left\{ [H_P]_N - [A]_N [H_{uv}]_N^{-1} [B]_N \right\} \quad (5.51)$$

This transformation converts the original 5X5 nodal admittance matrix into a 3X3 admittance matrix relating the node voltages at the CC-II terminals Y, X and Z to their correspondent input currents. The admittance matrix of the complete class AB CC-II, $[Y_T]$, is the sum of $[Y_T]_N$ with $[Y_T]_P$. Finally, the matrix form in (5.32) results after the following transformation,

$$[Y_T] = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} \rightarrow [T_T] = \begin{bmatrix} y_{11} - \frac{y_{12}y_{21}}{y_{22}} & \frac{y_{12}}{y_{22}} & y_{13} - \frac{y_{12}y_{23}}{y_{22}} \\ -\frac{y_{21}}{y_{22}} & \frac{1}{y_{22}} & -\frac{y_{23}}{y_{22}} \\ y_{31} - \frac{y_{32}y_{21}}{y_{22}} & \frac{y_{32}}{y_{22}} & y_{33} - \frac{y_{32}y_{23}}{y_{22}} \end{bmatrix} \quad (5.52)$$

where $[T_T]$ represents the CC-II transfer matrix according to definition (5.32).

Solving (5.52) explicitly result in a set of complex expressions for each transfer function within the CC-II matrix. These complex expressions are difficult to analyze due to the complexity of the involved factors. In general these expressions do not allow to draw useful information for circuit design. A better course of procedure is to use to numerical simulation of (5.52). Since equations (5.49), (5.51), (5.52) and Table 5.1 were developed using a normalized approach, numerical simulation can disclose important facts of the circuit's frequency response. In order to accomplish this task it is necessary to complete the transfer function definitions in (5.32). Equation (5.32) shows only the most relevant CC-II's transfer functions, which are: the current gain $a_i = I_Z/I_X$, the voltage gain $a_v = V_X/V_Y$ and the driving point impedances (admittances) $z_x = V_X/I_X$, $y_y = V_Y/I_Y$ and $y_z = V_Z/I_Z$. The remaining elements are ideally 0 for a CC-II. Unfortunately this is only approximated

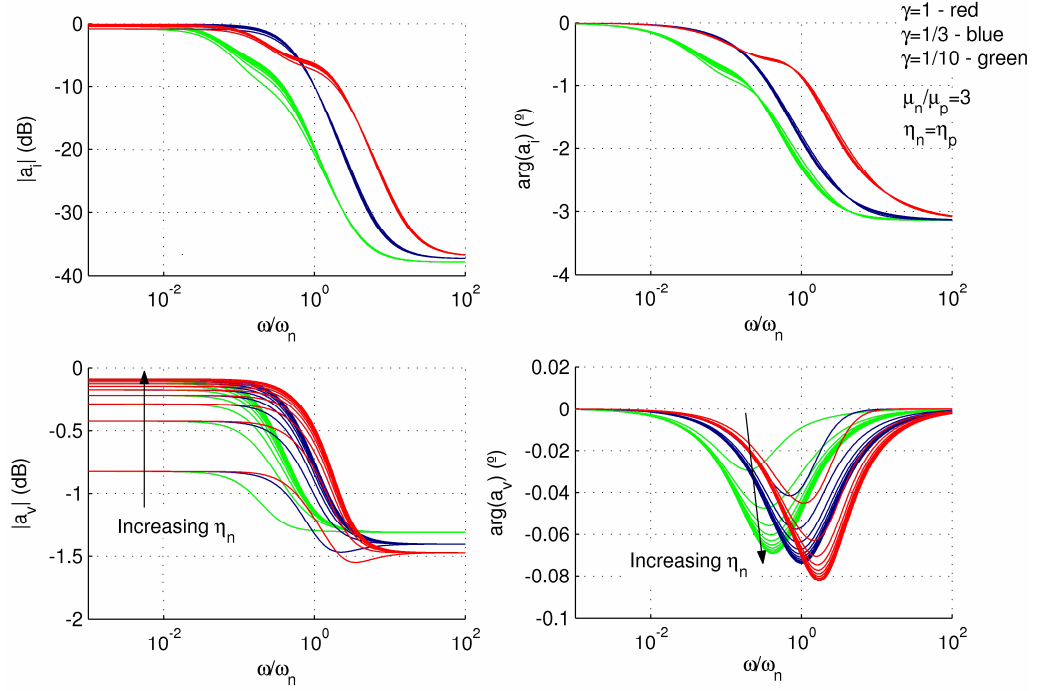


Figure 5.25 – CC-II's voltage gain and current gain frequency response (magnitude and phase).

by real circuits. For the present case, equation (5.32) should be replaced by,

$$\begin{bmatrix} I_Y \\ V_X \\ I_Z \end{bmatrix} = \begin{bmatrix} y_y & a_{xy} & y_{zy} \\ a_v & z_x & a_{zx} \\ y_{yz} & a_i & y_z \end{bmatrix} \begin{bmatrix} V_Y \\ I_X \\ V_Z \end{bmatrix} \quad (5.53)$$

where, a_{xy} , a_{zx} , y_{zy} and y_{yz} represent non-ideal transfer functions relating, I_Y to I_X , V_X to V_Z , I_Y to V_Z and I_Z to V_Y . The next step is to sweep the various normalized factors, on a one by one basis, and to observe the effect on the frequency response plots.

Figure 5.25 shows the frequency response plots, both magnitude and phase for several values of η_n and γ , assuming that $\eta_n = \eta_p$ (representing the case of NMOS and PMOS transistors having equal g_m/g_o ratios) and $\eta_c = 5$ (see Figure 5.24). The curves represented on Figure 5.25 were organized according to the value of γ , three different values, with three different colors (1, 1/3 and 1/10, represented by the colors red, blue and green, respectively). Within each color set, the parameter η_n assumed a variation from 10 to 100. As Figure 5.25 shows, the current gain is essentially immune to η_n variations. The same is verified for the voltage gain, since both its magnitude and phase variations with η_n are negligible; the magnitude does not even shows 3dB of variation. This kind of behavior is also observed in practical implementations of this circuit [288, 289, 292].

On the other hand, the current gain does not have such an ideal bandwidth characteristic. The presence of at least two poles is observed through the phase variation plots, corresponding to a total π radians phase rotation. The level off of the current gain curves near the -40dB together with the

step like behavior near the cut-off region, suggest the presence of more complex pole-zero dynamics on the current gain transfer function. Both current and voltage gain show a strong dependence on the parameter γ . To understand the meaning of these effects it is necessary to study the relation between γ and ρ , since these parameters are intimately connected. Since the transistors are biased with the same current and L_D is assumed equal for both PMOS and NMOS devices (implying equal η_c), the transconductance ratio is defined by,

$$\gamma = \frac{g_{m_p}}{g_{m_n}} = \sqrt{\frac{\mu_p}{\mu_n} \frac{W_p}{W_n}} \quad (5.54)$$

Accordingly, the ratio between transition frequencies becomes,

$$\rho = \frac{\omega_n}{\omega_p} = \frac{\tau_p}{\tau_n} = \frac{g_{m_n}}{g_{m_p}} \frac{C_{gd_p}}{C_{gd_n}} = \frac{1}{\gamma} \frac{W_p}{W_n} \quad (5.55)$$

Solving (5.54) for W_p/W_n and substituting the result on (5.55), delivers the relation,

$$\rho = \gamma \frac{\mu_n}{\mu_p} \quad (5.56)$$

Equation (5.56) shows that ρ and γ have a linear relation proportional to the mobility ratio between NMOS and PMOS transistors. It is possible to define either γ or ρ with any desired value, and equation (5.56) establishes the value of the remaining parameter. A direct implication of (5.56) is that the frequency matching condition, $\rho=1$, is constrained by technological parameters, namely the mobility ratio. The mobility ratio is not fixed for all CMOS process. In fact, is not even fixed within a single process. The mobility is a measure of the minority carrier's velocity dependence on the electrical field, intimately related with the doping profiles [253, 260, 262, 274]. CMOS technologies exhibit a phenomenon known as velocity saturation. Under low field conditions, the velocity of the minority carriers in a sample of doped silicon (P or N) is mainly proportional to the applied electrical field. This proportionality factor is the mobility factor. As the magnitude of the electrical field increases, this proportionality is destroyed and the minority carriers attain a saturation velocity. CMOS technology providers usually specify the effective mobility, as the mobility factor measured under low electric field conditions. For the purpose of this work, the effective mobility is used as a means to provide the mobility ratio, without entering into much detail velocity saturation effects and modeling. Nevertheless, it is important to recall that velocity saturation is one of the major limiting effects that can constrain both the dynamical and static operation of the MOSFET. It is especially relevant for short channel devices like devices currently in use.

Figure 5.26 shows the mobility ratio for different CMOS technological providers and different

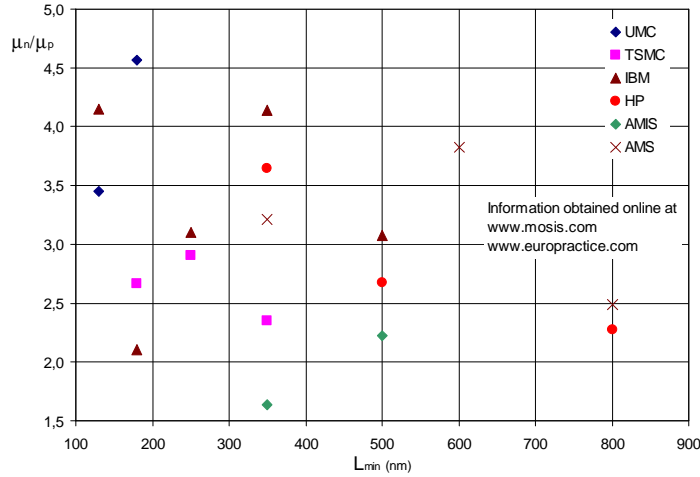


Figure 5.26 – Mobility ratio for different technological nodes and different foundries.

technological nodes. It is apparent that the mobility ratio does not follow any specific tendency, but is rather a result of other optimizing ends. It is also observable that the mobility ratio may display very different values for different technological nodes, even within the same provider.

Figure 5.25 shows that for $\mu_n/\mu_p=3$ the frequency response seems to approach a maximally flat frequency response when $\gamma=1/3$, that is, when $\rho=1$. The other situations, $\gamma=1/10$ and $\gamma=1$ (corresponding to matched transconductance) exhibit less desired frequency response characteristics. There is an inherent bandwidth loss and the frequency response has not maximally flat characteristics. This can be easily explained: since $\rho \neq 1$ implies PMOS and NMOS transistors with different transition frequencies, imperfect pole-zero cancellation behaviors may occur, producing the step like frequency response observed on Figure 5.25. The general consequence is bandwidth loss due to the fact that slower devices impose an early cut-off frequency. This effect does not occur when $\rho=1$ since the transistors have then similar transition frequencies.

Figure 5.27 and Figure 5.28 (see next page) show other CC-II transfer functions. These transfer functions (except for the reverse voltage and current gains, a_{zx} and a_{xy} , respectively) are represented in their normalized version, taking as normalizing factor the admittance Y_{norm} , given by the sum of Y_0^N with Y_0^P (see Table 5.1). Figure 5.27 depicts the normalized terminal driving point impedance, z_x and the driving point admittances, y_y and y_z . Both z_x and y_y are not very susceptible to η_n or η_p variations. They are however strongly dependent on γ , as γ increases the magnitude curves of z_x and y_y shift to higher frequencies. On the other hand, y_z depends strongly on all the parameters; η_n and η_p produces significant changes on magnitude, and γ introduces the same frequency shift effect as above. From the magnitude and phase plots of these driving point impedances (admittances) it is also apparent that:

- i) z_x can be roughly represented by a parallel association of a capacitance and a resistance, since

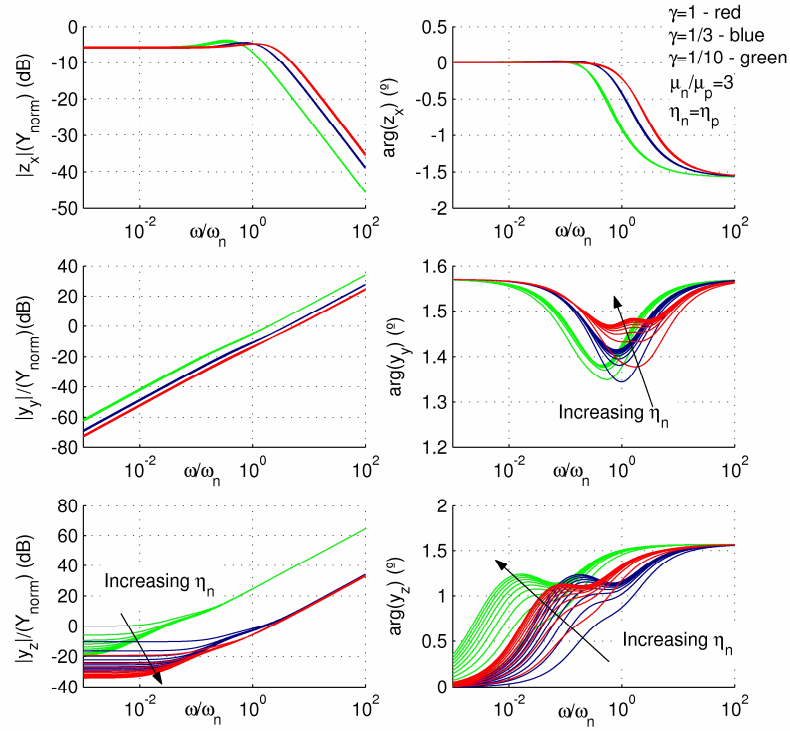


Figure 5.27 – Terminal driving point impedances (admittances); both magnitude and phase.

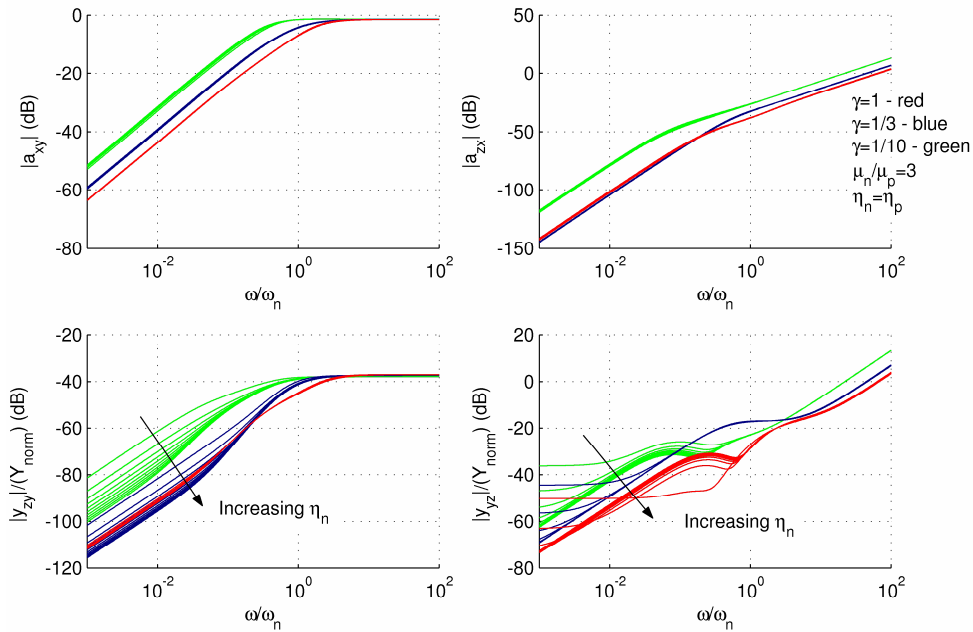


Figure 5.28 – Non-ideal transfer functions (reverse gains and reverse transfer admittances); magnitude only.

it provides a high frequency roll-off of -20dB/dec and a total phase shift of $\pi/2$ radians (ignoring the small overshoot effect near the cutoff region);

ii) y_y is mainly capacitive, since it grows almost linearly with frequency while the phase is almost constant (near $\pi/2$ as it must be for a ideal capacitor); and,

iii) y_z can also be represented by a capacitance in parallel with a resistance, since the same

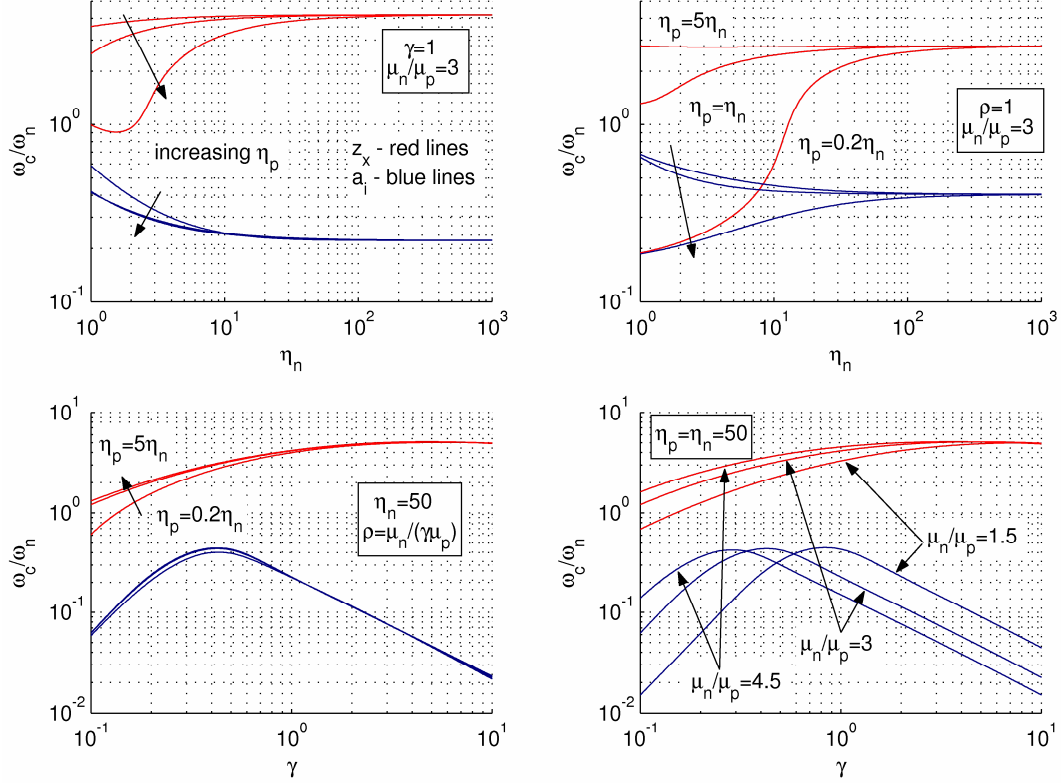


Figure 5.29 – CC-II's bandwidth performance evaluation

characteristics as for the case of z_x are also evident.

Figure 5.28 depicts the frequency response (only magnitude plots) of the non-ideal transfer functions. Except for the reverse current gain a_{xy} , the other transfer functions (the reverse voltage gain a_{zx} and the transfer admittances y_{zy} and y_{yz}) have negligible magnitudes for normalized frequencies (ω/ω_n) below unity. On the other hand a_{xy} approaches unity near $\omega/\omega_n=1$ and is negligible for smaller frequencies. It is common practice to disregard the contribution of these four transfer functions. Figure 5.28 shows that this simplification does not produce severe errors, for frequencies within the frequency range of interest. Thus it is possible to assume that, for positive CC-II employing translinear loops, the non-ideal transfer functions have approximately zero value.

5.4.6 Design Guidelines for Large Bandwidth Translinear CC-IIs

Last section presents a modeling strategy based on normalized transfer functions to evaluate the frequency behavior of translinear CC-IIs. Using this approach, the ultimate bandwidth performance of these CC-IIs is indexed to the transition frequency of its NMOS devices. Applying the model synthesized on equations (5.49), (5.51), (5.52) and Table 5.1, furnishes a suitable method to evaluate the cut-off bandwidth of such circuits. Figure 5.29 presents a study of the normalized cut-off frequency dependence on several design parameters. From the previous discussion, it was established that the most important transfer functions on a translinear CC-II are the current gain, a_i

and the driving point impedance at port X, z_x . These two transfer functions are, from the point of view of CMD's conception, the most stringent in what regards bandwidth; z_x , because it promotes current division at the input port³⁵, and a_i because is the main element used to convey current from the input to the output ports.

The first plots on the top of Figure 5.29 present the normalized cut-off variation against η_n and η_p assuming two cases, matched transconductance, $\gamma=1$, and matched transition frequencies, $\rho=1$. Three facts deserve attention. First, the normalized cut-off frequencies of both a_i and z_x are not affected by η_n nor η_p for large values of these parameters, that is, the CC-II's cut-off frequency does not depend on the output conductance of its transistors for large values of η_n and η_p . Second, when η_n and η_p have small values, both cut-off frequencies exhibit large variations. In particular the current gain's cut-off frequency may exhibit large values for smaller η_n and η_p . However, values of η_n and η_p near unity are only a theoretical abstraction; real MOSFET transistors usually have transconductance values larger than the output conductance³⁶. Third, the condition $\rho=1$ shows superior frequency performance than $\gamma=1$, achieving a cut-off frequency near $0.4\omega_n$ (for large values of η_n and η_p) compared to $0.2\omega_n$ for the last case.

The bottom plots on Figure 5.29 reveal important frequency response optimization guidelines. Instead of fixing γ (or ρ), these results presents the cut-off frequency dependence on parameter γ . One thing that is readily apparent is that there is an optimum value for γ corresponding to the maximum possible cut-off frequency. This maximum behavior is present for both a_i and z_x transfer functions (for the case of z_x , the maximum occurs for larger values of γ and exhibits larger frequency values, near $5\omega_n$). These results, especially the bottom right plot, clearly demonstrate the dependence of the maximum cut-off frequency on the mobility ratio. It was previously discussed that the frequency matching condition $\rho=1$ would provide the best possible cut-off frequency. The results shown so far have supported this conclusion, without furnishing clear evidence. The cut-off frequency plot on the bottom right corner of Figure 5.29 depicts this occurrence. This plot shows that a_i 's cut-off frequency has a maximum and this maximum shifts from left to right when the mobility ratio (μ_n/μ_p) decreases. It was also investigated the conditions on γ that provide the maximum cut-off, showing that for such γ values the condition $\rho=1$ was met. Finally, it is possible to conclude that in order to achieve the maximum cut-off frequency, the CC-II has to be designed

³⁵ It is also possible to have current division at the output port, but its effect is not as limiting as it is in the input side. This is due to the presence of a large capacitance photo detector at the input.

³⁶ Obviously, it is possible to force artificially smaller values of η_n and η_p if that proves profitable.

with transistors satisfying the condition $\rho=1$, or equivalently, $W_n/W_p=\mu_n/\mu_p$. This would allow a cut-off frequency near $0.4\omega_n$, slightly dependent on the output conductance of the transistors in use.

From the CMD design perspective, it is also important to cope with input impedance seen at port X. As it was previously discussed, the net impedance at port X is inversely proportional to $Y_{norm}=Y_0^N+Y_0^P$. Looking to Figure 5.27, the proportional factor is near 0.5. Once Y_0^N and Y_0^P are known, it is possible to have an estimative of $|z_{xx}(0)|$ and the cut-off frequency.

5.5 Input Impedance Reduction

There are several possibilities to reduce (or enhance) a specific circuit driving point impedance. Once the appropriate factors that establish the required impedance are correctly identified this can be done at the circuit level. Usually, working at the circuit level requires changing circuit layout and area or even changing bias conditions. When these changes lead to area or power inefficient solutions or even unfeasible designs (as is usually the case of impedance reduction), better strategies should be applied. An efficient strategy is to apply adequate feedback configurations to improve the design. Concerning impedance shaping, two feedback approaches can be employed to reduce impedance level: i) negative feedback comprising parallel associations of the feedback network with the required circuit port; or, ii) positive feedback comprising series combinations of the feedback network with the required circuit port. The first strategy is the most widely used since it leads to stable and well behaved designs. Nevertheless, positive feedback may be eventually applied with the appropriate care to achieve the same impedance reduction effect.

5.5.1 Available Techniques

As previously discussed, CC-II's can be used to implement the CMD concept (see Figure 5.11 and accompanying discussion). On such applications, the current transfer occur between the X input port and Z output port of the CC-II. The Y port can be used for other purposes, such as DC bias control. The important features that the CC-II has to satisfy are high bandwidth current gain and small input impedance at port X. In this sense, any available technique that can realize both of these quantities is welcomed.

It is possible to reduce the X input impedance using feedback configurations. For this purpose, it is a great advantage to explore the X port as both a current input (feeding the output current replica) and a voltage output (when controlled by the voltage input at port Y). This dual character, combined with the availability of the Y terminal in a CMD configuration, allows several feedback schemes. Figure 5.30 presents a survey of feedback schemes used for this particular purpose. These schemes comprise both positive and negative feedback configurations.

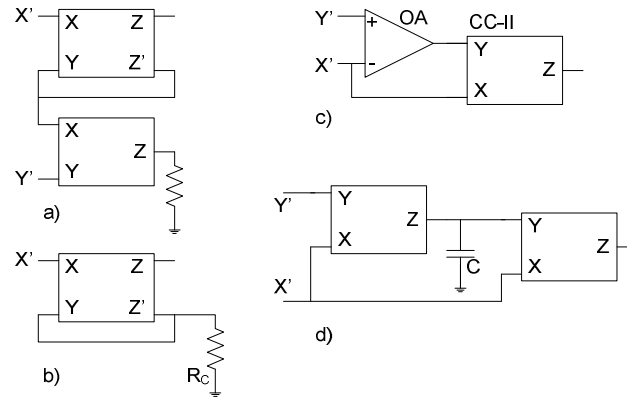


Figure 5.30 – Feedback configurations to reduce the X input impedance.

Figure 5.30a and b, depicts two variants of the same positive feedback configuration. On Figure 5.30a an auxiliary CC-II is used to provide an available Y terminal and also matched compensation impedance (as R_C in Figure 5.30b). Figure 5.30b represents a simplified version of Figure 5.30a where the Y terminal is not required³⁷. The configuration of Figure 5.30a was originally issued by Fabre and Barthelemy in [226] as an improved version of the CC-II. Later on, Palumbo and Pennisi presented the feedback approach version of Figure 5.30b in [233]. Looking to Figure 5.30b it is easily seen that the X port appears in a series connection with the feedback network (comprising R_C , an auxiliary Z port and Y port). Such series feedback configurations are generally used for enhancing terminal impedances. However, due to the positive character of the feedback (due to the polarity of the current gain seen at the auxiliary Z output), the effect on the X input impedance is exactly the opposite. In the past Nero Alves *et al* used variants of this scheme to produce a differential current buffer with reduced differential input impedance [287-289, 291, 292].

The remaining two configurations (Figure 5.30c and d) are negative feedback configurations. The first one on Figure 5.30c) is the traditional unitary voltage feedback configuration using an operational amplifier (OA). The first appearance of this technique dates back to the early work on feedback amplifiers by Black and Blackman [75, 78]. For this case the X terminal is viewed as a voltage output and the feedback loop appears connected in a parallel fashion. The resulting output impedance is reduced by an amount proportional to the OA's voltage gain.

Finally, the scheme of Figure 5.30d represents another negative feedback configuration. The X port is for this case a current input terminal, with the feedback acting in parallel with it. The net result translates in reduction of the input impedance seen at port X. The feedback loop comprises the two CC-IIs and the capacitance C. Ideally, the return ratio is very large for small signal frequencies, thus allowing significant reductions of the input impedance. In practice, y_z and y_y limit the action of

³⁷ It is possible to establish the DC bias condition using R_C connected to an adequate voltage reference.

the capacitor posing severe limitations on the effectiveness of this scheme. This technique was originally introduced by Arbel in [230] together with a dual-loop feedback variant of the same concept.

Dual-loop configurations may reduce even more the input impedance. However, these dual-loop possibilities result in severe design limitations, due to increased complexity, increased power, increased layout area and reduced stability margins. [230] explored the usage of dual-loop feedback employing the concepts of Figure 5.30c and d. The OA would in this sense compensate for the non-idealities of the CC-IIs. Another possibility explored by the Nero Alves *et al* is to combine positive and negative feedback types on the same scheme [295]. For this purpose, the schemes of Figure 5.30a and c were combined, allowing an effective control of the input impedance. The following discussion is centered on the basic single loop configurations of Figure 5.30, covering dual-loop configurations with less detail, due to their complexity and inherent limited application.

5.5.2 Positive Feedback

As it was previously disclosed the important figures that the CMD must satisfy are high bandwidth current gain and small input impedance. Rosenstark method and Blackman theorem (see

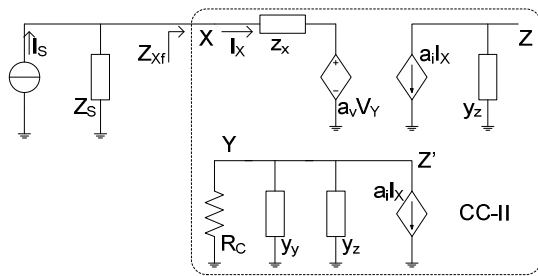


Figure 5.31 – Positive feedback scheme: small signal equivalent circuit.

appendix C 2) are the preferred tools to analyze the impact of feedback on both current gain and input impedance. Starting with the simplest feedback scheme of Figure 5.30b), comprising one CC-II and a feedback resistance R_C , the return ratio can be found using the small signal equivalent circuit of Figure 5.31. In this equivalent circuit, Z_S represents the source impedance (for the CMD case, Z_S can be represented by a single capacitance).

According to Blackman's theorem (appendix C 2), the driving point impedance seen at the input port of the CMD, z_{xf} , can be evaluated using a three step procedure:

- i) First, the driving point impedance is measured in open-loop configuration, z_{xol} (killing the feedback).
- ii) Second, the return ratio of the circuit is evaluated in two special situations: with the port under investigation shorted, $R(0)$, and opened $R(\infty)$.
- iii) Finally, the closed-loop driving point impedance is given by,

$$z_{xf} = z_{xol} \frac{1 + R(0)}{1 + R(\infty)} \quad (5.57)$$

$R(0)$ and $R(\infty)$ can be evaluated in a single test if a test impedance is connected to the port under

investigation. For the circuit of Figure 5.31, this test impedance can be Z_S . Then $R(0)$ and $R(\infty)$ are the limits of $R(Z_S)$ when Z_S tends to 0 (shorted condition) and ∞ (opened condition).

The return ratio $R(Z_S)$ is also required for the asymptotic gain model. Choosing I_X for both control and output variables and a_i as control parameter (the current gain of the CCCS associated with the auxiliary Z' port), the return ratio is by definition $R(Z_S) = -I_X$ when the CCCS is replaced by an independent current source of value a_i . That is,

$$R(Z_S) = -\frac{a_i a_v R_C}{(z_x + Z_S)(1 + R_C(y_y + y_z))} \quad (5.58)$$

The direct gain term, G_o , is defined as the ratio of I_X to I_S when $a_i = 0$. Making $a_i = 0$ on the auxiliary Z terminal, results in $V_Y = 0$ and I_X is given as a current division between Z_S and z_x , that is,

$$G_o = \left. \frac{I_X}{I_S} \right|_{a_i=0} = \frac{Z_S}{Z_S + z_x} \quad (5.59)$$

Similarly, the asymptotic gain term is given by the ratio between of I_X to I_S when a_i goes to infinity. Since V_Y cannot have an infinite value, this means that I_X is set to 0 under this condition, thus $G_\infty = 0$. Finally, the open-loop input impedance, z_{xol} , is the input impedance seen at port X when $a_i = 0$, that is $z_{xol} = z_x$.

Now, using (5.58) with $Z_S = 0$ and $Z_S = \infty$ together with Blackman theorem results on the closed-loop input impedance, z_{xf} , given by,

$$z_{xf} = z_{xol} \frac{1 + R(0)}{1 + R(\infty)} = z_x - \frac{a_i a_v R_C}{1 + R_C(y_y + y_z)} \quad (5.60)$$

Equation (5.60) Shows that the effect of R_C on z_{xf} is to reduce z_x by an amount approximately proportional to R_C (when y_z and y_y are negligible). According to (5.60) it is even possible to produce negative impedances using this technique. However, such possibility is of limited application from the CMD conception point of view.

Finally, using G_o , G_∞ and $R(Z_S)$ together with the Rosenstark asymptotic gain model, results on the closed-loop current gain (referenced to the output Z) given by,

$$\begin{aligned} a_{if} &= \frac{I_Z}{I_S} = a_i \frac{I_X}{I_S} = a_i \frac{G_\infty R(Z_S) + G_o}{1 + R(Z_S)} \\ &= a_i \frac{Z_S}{Z_S + z_x - \frac{a_i a_v R_C}{1 + R_C(y_y + y_z)}} = a_i \frac{Z_S}{Z_S + z_{xf}} \end{aligned} \quad (5.61)$$

Equation shows that the current gain is also improved by the feedback action, since reducing z_{xf} results in improved current transfer between the current source and the CMD.

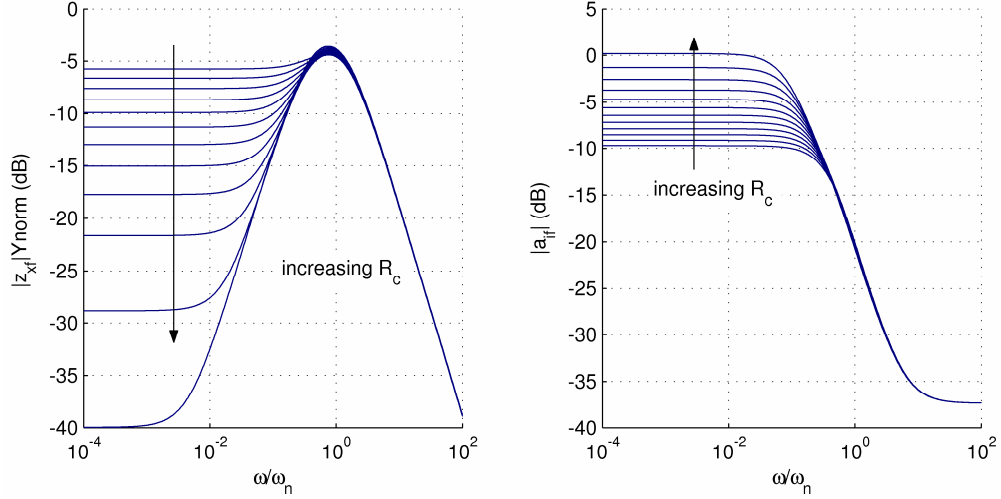


Figure 5.32 – Feedback effect on input impedance and current gain, positive feedback scheme.

The frequency behavior of equations (5.60) and (5.61) can be analyzed using the CC-II normalized model of the previous section, Figure 5.32 depicts one such analysis. The CC-II model parameters were $\rho=1$, $\eta_c=5$, $\eta_p=\eta_n=50$ and $\mu_n/\mu_p=3$. The source impedance was assumed to be a pure resistance with a value set to $Z_S=|z_x(0)|$, which represents a small and pessimistic value but serves better the purpose of behavior evaluation. R_C was also normalized with reference to $|z_x(0)|$, however its effect on feedback follows straightforward, since the normalization is taken only on a proportional basis. Y_{norm} in Figure 5.32 represents the normalizing impedance, given by $Y_o^N+Y_o^P$, as in section 5.4.5. Figure 5.32 shows that increasing R_C improves both z_{xf} (which becomes smaller) and a_{if} (which approaches unity). The high frequency behavior is mainly limited by the CC-II frequency response (a_i on a_{if} and z_x on z_{xf}). One important observation is that z_{xf} does not maintain its reduced value for all the available z_x bandwidth. This is due to the fact that a_i and z_x have different cut-off frequencies, with a_i cutting earlier than z_x ³⁸, thus reducing the feedback action for frequencies above the cut-off frequency of the current gain. Associated to this effect is also the presence of non-ideal terminal impedances y_y and y_z , which together with R_C contribute to restrict the feedback action.

5.5.3 Negative Feedback with OA

Figure 5.33 depicts the small signal equivalent circuit of the negative feedback configuration of Figure 5.30c. The input Y' of the composite circuit is not used for signal processing ends, nevertheless it can serve as DC bias establishing point. Thus, for small signal analysis it can be assumed that Y' is a grounded terminal. It is assumed that the OA is adequately represented by a

³⁸ Note that a_v is essentially fixed over all the important frequency range as shown previously on Figure 5.25.

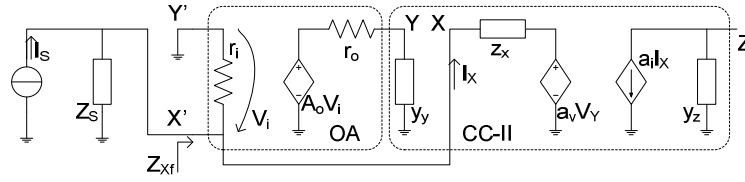


Figure 5.33 – Negative feedback with OA scheme: small signal equivalent circuit.

differential input resistance r_i , a VCVS controlled by the input voltage V_i with voltage gain A_o in series with an output resistance r_o . Under these assumptions the closed-loop input impedance and current gain are determined in a similar fashion as in the previous case (that is, using Blackman Theorem and the asymptotic gain model). Selecting V_i as the control variable, A_o as the control parameter and I_X as the output variable, the return ratio $R(Z_S)$ is given by $-V_i$ when the VCVS is substituted by an independent voltage source with value A_o , resulting in,

$$R(Z_S) = \frac{a_v A_o}{1 + r_o y_y} \frac{Z_S'}{Z_S' + z_x} \quad (5.62)$$

where Z_S' represents the parallel association of Z_S and r_i . The direct gain term is defined as the ratio of I_X to I_S when $A_o = 0$. Making $A_o = 0$ results in $V_Y = 0$ and I_X given as a current division between Z_S' and z_x , that is,

$$G_o = \frac{I_X}{I_S} \Big|_{A_o=0} = \frac{Z_S'}{Z_S' + z_x} \quad (5.63)$$

Similarly, the asymptotic gain term is given by the ratio between of I_X to I_S when A_o goes to infinity. Since V_Y cannot have an infinite value, this means that V_i must be set to 0 under this condition, thus $I_X = I_S$ and $G_\infty = 1$. Finally, the open-loop input impedance, z_{xol} , is the input impedance seen at the input terminal X' when $A_o = 0$, that is $z_{xol} = z_x // r_i$. Since $R(0) = 0$, the closed-loop input impedance z_{xf} is given by,

$$z_{xf} = \frac{z_{xol}}{1 + R(\infty)} = \frac{z_x}{1 + \frac{a_v A_o}{1 + r_o y_y} + \frac{z_x}{r_i}} \quad (5.64)$$

Equation (5.64) shows that large gains OAs produces small impedance levels. However, large gains OAs have restricted bandwidths, thus implying less effective feedback action. Finally the closed-loop current gain follows directly from the asymptotic gain formula,

$$a_{if} = a_i \left(\frac{G_\infty R(Z_S) + G_o}{1 + R(Z_S)} \right) = a_i \frac{Z_S'}{Z_S' + z_{xf}^\infty} \quad (5.65)$$

where z_{xf}^∞ represents the limit of z_{xf} when r_i becomes infinity. As in the previous case, the effect of the feedback action is to reduce current division between Z_S and the input of the composite CC-II, thus allowing the closed-loop current gain to approach the ideal current gain of the CC-II.

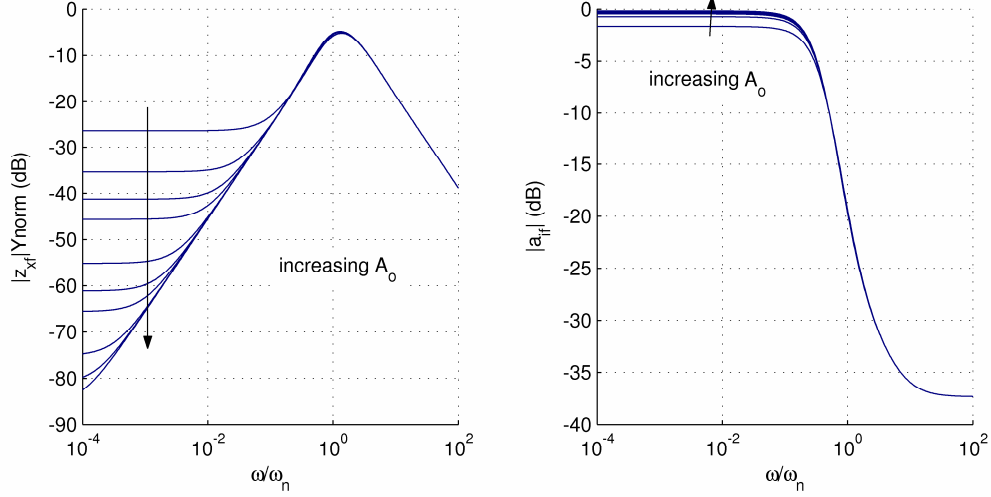


Figure 5.34 - Feedback effect on input impedance and current gain, negative feedback with OA.

In the above analysis it was assumed that the OA have an ideal voltage gain A_o independent of the frequency. In practice this is not a good approximation. In order to have a closer look over the frequency behavior of equations (5.64) and (5.65), the OA must have an adequate frequency response model. For that purpose, it was assumed that the OA can be modeled accurately by a first order low pass transfer function, corresponding to a unity gain frequency equal to ω_h (the transition frequency of the NMOS transistors in the CC-II). Under this assumption, the gain and cut-off frequency of the OA are related by $A_o\omega_c = \omega_h$. Using a CC-II with the same parameter set as in the previous case, the resulting closed-loop frequency response is depicted on Figure 5.34. It is readily apparent that this technique can exhibit superior performance than the previous one in terms of impedance reduction. In fact, the input impedance can attain very small values, if the OA as a very large voltage gain. However, the frequency range on which the feedback is effective is generally less than on the previous case. This is a consequence of the gain-bandwidth limitations in OAs. In common practice, OAs may exhibit similar non-ideal dynamics; even though the first order dynamics may result from design choices, the general bandwidth limitations predict that gain and bandwidth have inverse dependencies.

5.5.4 Negative Feedback with Two CC-IIs

Figure 5.35 depicts the small signal equivalent circuit of the negative feedback configuration of Figure 5.30d. As in the previous case, the Y' terminal is assumed to be a grounded terminal for small signal analysis. The analysis proceeds in a similar fashion, as on the previous cases. Selecting I_1 as the control variable, a_i as the control parameter (from the first CC-II on the left of Figure 5.35) and I_2 as the output variable, the return ratio $R(Z_S)$ is given by $-I_1$ when the CCCS is substituted by an independent current source with value a_i , resulting in,

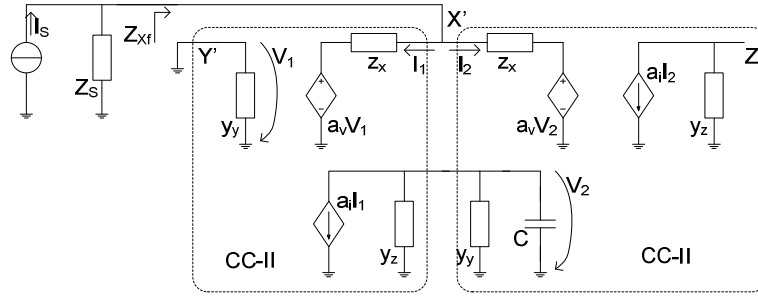


Figure 5.35 - Negative feedback with two CC-IIs scheme: small signal equivalent circuit.

$$R(Z_S) = \frac{a_i a_v}{z_x (y_y + y_z + sC)} \frac{Z_S'}{Z_S' + z_x} \quad (5.66)$$

where Z_S' represents the parallel association of Z_S and z_x . The direct gain term is defined as the ratio of I_2 to I_S when $a_i=0$. Making $a_i=0$ results in $V_2=0$ and G_o is given by,

$$G_o = \left. \frac{I_x}{I_s} \right|_{a_i=0} = \frac{Z_S'}{Z_S' + z_x} \quad (5.67)$$

Similarly, the asymptotic gain term is given by the ratio between of I_2 to I_S when a_i goes to infinity. Since V_2 can not have an infinite value, this means that I_1 must be set to 0 under this condition, thus $I_2=I_S$ and $G_\infty=1$. Finally, the open-loop input impedance, z_{xol} , is the input impedance seen at the input terminal X' when $a_i=0$, that is $z_{xol}=z_x/2$. Since $R(0)=0$, the closed-loop input impedance z_{xf} is given by,

$$z_{xf} = \frac{z_{xol}}{1 + R(\infty)} = \frac{z_x/2}{1 + \frac{a_v a_i}{2z_x (y_y + y_z + sC)}} \quad (5.68)$$

Equation (5.68) shows that C can be used to reduce the resistance level of z_{xf} . Assuming that y_y and y_z are negligible, equation shows that for small frequencies z_{xf} is 0. In common practical applications, the presence of non-ideal values of y_y and y_z limits the effectiveness of the method. This is because the resistance level of y_z+y_y establishes a lower bound for the values of C . Finally the closed-loop current gain follows directly from the asymptotic gain formula,

$$a_{if} = a_i \left(\frac{G_\infty R(Z_S) + G_o}{1 + R(Z_S)} \right) = a_i \frac{\frac{1}{2} + \frac{a_v a_i}{2z_x (y_y + y_z + sC)}}{1 + \frac{a_v a_i}{2z_x (y_y + y_z + sC)}} \quad (5.69)$$

Equation (5.69) shows that a_{if} approaches unity as $R(Z_S)$ tends towards infinity and reaches 0.5 when $R(Z_S)$ approaches 0.

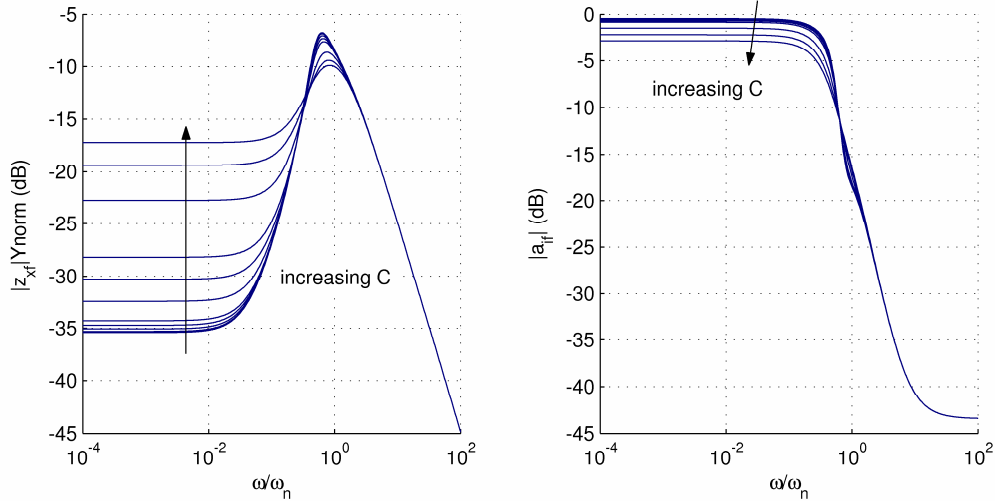


Figure 5.36 - Feedback effect on input impedance and current gain, negative feedback with two CC-IIs.

Figure 5.36 shows the frequency behavior of equations (5.68) and (5.69), assuming the same CC-II model as before for both CC-IIs of this configuration. The effect of y_y and y_z is clearly evident on $|z_{xf}|$; for small values of C , the resistance level of z_{xf} is essentially fixed by the CC-II itself. On the other hand, large values of C are not useful, since z_{xf} approaches z_{xol} . It is also apparent, from Figure 5.36, that the frequency range for which $|z_{xf}|$ maintains a constant resistance level is larger than on the previous methods. The effect of feedback on the current gain is essentially the same as in the previous cases; as z_{xf} is reduced the current gain approaches unity and the high frequency roll-off is dictated by the CC-II internal current gain.

5.5.5 Performance Comparison

Figure 5.37 (see next page) presents the results of a performance comparison of the three methods discussed on the previous sections. Since the main aim of employing feedback to CC-II based CMD design techniques is primarily to reduce or control the input impedance, it is important to have a common basis to evaluate the performance of the proposed methods. A set of two parameters were adopted for this purpose, both related to the input impedance under feedback. The first parameter measures the normalized frequency range (against ω_n), on which the input impedance has flat amplitude response. The second parameter measures the height of the high frequency peak (as discussed in relation to Figure 5.32, Figure 5.34 and Figure 5.36); this is also a measure of the impedance reduction due to feedback, since the peak amplitude is mainly due to the open-loop input impedance. The comparison study uses a generic control variable u , in order to have a common control parameter. This generic variable represents the resistance, R_C , of the first method, the voltage gain, A_o , of the OA of the second method and the capacitance, C , of the third method. Since these parameters are all taken in their normalized form ($R_{Cn} = R_C / |z_x(0)|$), A_o itself and

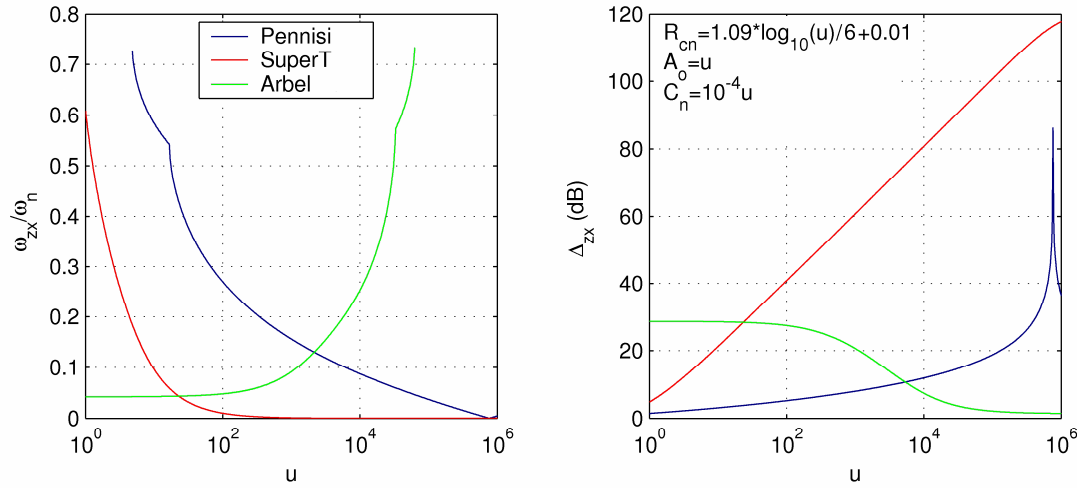


Figure 5.37 – Performance comparison.

$C_n = \omega_n C / Y_{norm}$, respectively), they are all non dimensional parameters, thus, the same generic variable u can be adequately adopted to represent each of these normalized parameters (R_{Cn} , A_o and C_n). The transformations used in this study were,

$$\begin{aligned} R_{Cn} &= 1.09 \log_{10}(u) / 6 + 0.01 \\ A_o &= u \\ C_n &= 10^{-4} u \end{aligned} \quad (5.70)$$

Both A_o and C_n follow a logarithmic distribution ranging through six decades of variation. R_{Cn} on the other hand follows a linear distribution, ranging from 0.01 (nearly open-loop) to 1.1 (slightly above $|z_x(0)|$). Figure 5.37 shows that the most efficient method to reduce the impedance level is the second method consisting in a negative feedback configuration with an OA (represented in Figure 5.37 by the red lines with the label “SuperT” recalling the super-transistor concept), requiring only very large gain OAs. On the other hand, this is also the method that presents worst bandwidth limitations, since there is an inherent tradeoff between impedance reduction (inversely proportional to the gain of the OA) and bandwidth. Since a CMD must have flat low impedance level input impedance, it is better to use the first method (positive feedback, represented in Figure 5.37 by the blue lines and the label “Pennisi” recalling its author). This is the method that presents the best compromise between impedance reduction and bandwidth; it is also the simplest and least resource demanding of all.

Figure 5.37 shows that using positive feedback may also lead to negative impedance conversion. The impedance reduction plot shows a peak value of more than $80dB\Omega$, near $u=10^6$ ($R_{Cn}=1.1$). According to (5.60) this peak value represents the 0Ω input impedance condition. Increasing further R_{Cn} shows that the input impedance changes to negative values with increasing magnitudes. Finally, Figure 5.37 shows that the negative feedback with two CC-II's method is the least appealing. This

method (green lines with the label “Abel” recalling its author) is restricted by the CC-II terminal admittances, evident on the flat behavior for small values of u . As u increases the impedance reduction decreases towards zero starting with values near $40dB\Omega$.

As a final comment, it is possible to say that the positive feedback configuration is the one that shows best performance when both impedance reduction and bandwidth represent important design considerations. It is also the best choice employing feedback in what concerns simplicity and resource demands.

5.5.6 Dual-Loop Positive-Negative Feedback

Sections 5.5.2, 5.5.3 and 5.5.4 presented three single loop feedback configurations able to reduce the X input impedance of a CC-II. It is possible to combine these configurations into more elaborate circuits comprising multiple feedback loops. Arbel reported for the first time a dual loop configuration consisting in the combination of the methods explored in sections 5.5.3 and 5.5.4 [230]. This dual loop configuration uses two negative feedback loops, one based on the configuration with two CC-IIs, and the other based on the OA. The advantage of this scheme was to combine the effects of both methods and add another controlling variable to the system dynamics. For this case, both a capacitance and the OA’s voltage gain can be used to control the X input impedance.

Nero Alves *et al*., reported another dual loop configuration comprising both positive and negative feedback loops, having similar complexity [295]. To this end, the methods of section 5.5.2

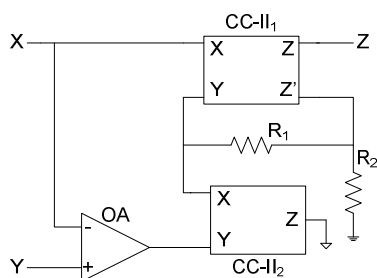


Figure 5.38 – Dual loop positive-negative feedback.

and 5.5.3 were combined, leading to the circuit of Figure 5.38. The two CC-IIs form the positive feedback loop; in fact this portion of the circuit is similar to the circuit of Figure 5.30a, except for the two resistances, R_1 and R_2 . The OA implements the negative feedback loop. In theory, it is possible to eliminate the two resistances, simply by shorting R_1 and removing R_2 . This would result in the extreme positive feedback case, where the X

input impedance of CC-II₁ is reduced ideally by its exact replica furnished by CC-II₂ (assuming CC-IIs with match characteristics). However, this is often difficult to synthesize in practical circuits due to stability restrictions. Resistances R_1 and R_2 eliminate this problem, providing an easy method to control the amount of positive feedback. The OA adds another form of control. According to section 5.5.3, the OA allows to reduce the X input impedance by a factor inversely proportional to its voltage gain. However, this also results in severe bandwidth penalties, since the bandwidth of the

OA is major limiting factor. Using the circuit of Figure 5.38 this problem can be reduced, since it is possible to use the positive feedback branch to compensate for small gain OAs.

It is possible to use both Rosenstark's asymptotic model and Blackman's theorem to perform the analysis of this circuit. However, it is simpler to use the results of sections 5.5.2 and 5.5.3 to accomplish the same task. Assuming that the two CC-IIs have identical characteristics and that the OA has voltage gain A_v and ideal input and output resistances (∞ and 0 respectively), the X input impedance can be found, including the effect of R_1 , R_2 and CC-II₂ on (5.60) and combining the result with (5.64). Then z_{xf} is given by

$$z_{xf} = z_x \frac{1 - K_2 a_i a_v}{1 + K_1 K_2 a_v^2 A_v} \quad (5.71)$$

Where K_1 and K_2 are proportionality factors defined as

$$K_1 = 1 + \frac{R_1}{R_2} (1 + y_z R_2)$$

$$K_2 = \left[(1 + y_y z_x) K_1 + \frac{z_x}{R_2} (1 + y_z R_2) \right]^{-1} \quad (5.72)$$

According to (5.72) K_1 is potentially larger than unity, while K_2 is in general smaller than unity. In (5.71) K_2 controls the amount positive feedback, while $K_1 K_2 A_v$ controls the negative feedback. In general $K_1 K_2$ is approximately unity, so A_v is the parameter that effectively controls the amount of negative feedback. As in the previous cases (sections 5.5.2, 5.5.3 and 5.5.4) it is possible to explore the effect of feedback in (5.71) using numerical simulation. Using the same CC-II normalized models as before gives the results of Figure 5.39. Figure 5.39 shows how R_1/R_2 and A_o (the magnitude of A_v) can be used to adjust the X input impedance characteristics. As before, two measures are observed, the normalized cut-off frequency, and the height of the high frequency peak, associated to z_{xf} . As can be seen, the two controlling parameters and A_o have opposite effects both on cut-off and impedance reduction. Increasing R_1/R_2 increases both cut-off frequency and the magnitude of z_{xf} . On the other hand, increasing A_o , reduces both cut-off frequency and the magnitude of z_{xf} . This suggests that A_o can be adjusted to establish the level of impedance reduction and the adequate cut-off characteristics. Then, the desired impedance is adjusted using R_1/R_2 .

There is also another parameter affecting the control dynamics of this dual loop configuration. The ratio z_x/R_2 has great impact on the overall dynamics. In general, it is desirable to have R_2 larger than z_x . However, a close inspection of Figure 5.39 shows that z_x/R_2 must have an optimum value, leading to the best possible performance.

Comparing the performance of this technique with the techniques reported in sections 5.5.2,

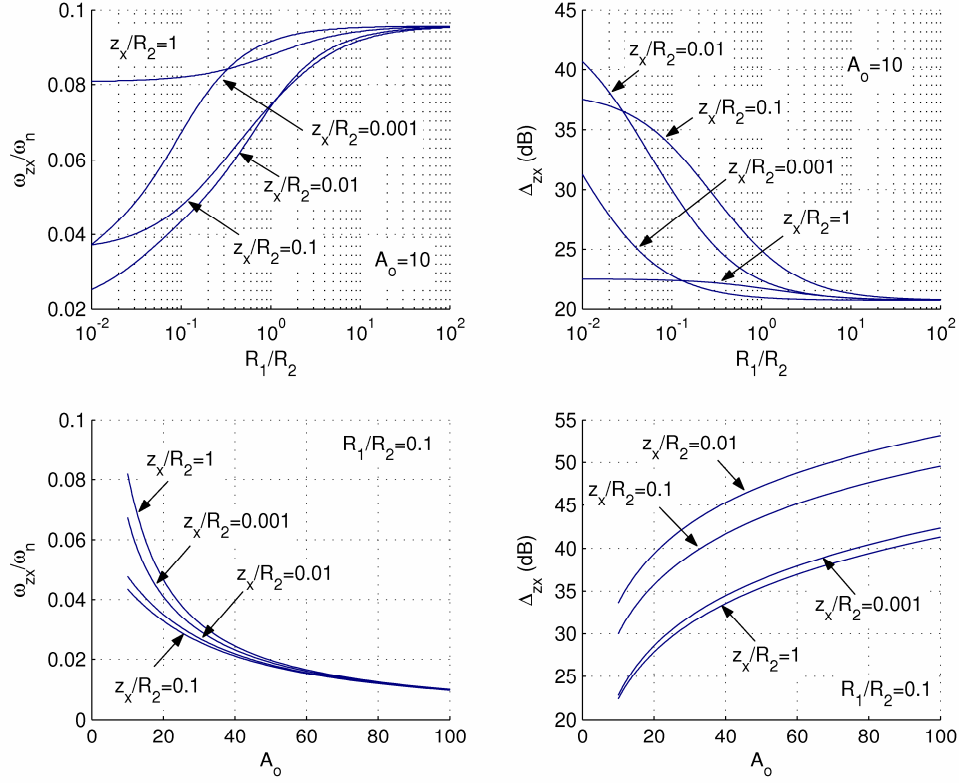


Figure 5.39 – Dual loop positive-negative feedback: effect of feedback on impedance.

5.5.3 and 5.5.4 it can be concluded that this technique has an impedance reduction effectiveness comparable to the method of section 5.5.3 (negative feedback with OA), with larger frequency response capabilities. The enhanced frequency response results from the trading between positive and negative feedback possible in this configuration. However, from a circuit design perspective this technique is not as competitive as the simple positive feedback configuration. The increased complexity may impair the usage of this configuration, since it generally results in larger power consumption, larger implementation areas, and complex control mechanisms.

5.6 Concluding Remarks

This chapter presented some preliminary considerations of current-mode design techniques, focused on bandwidth limitations. The application of current-mode design to the new concept of current matching (CMD) was also discussed. The concept of current matching is based on the existence of a composite active device able to convey large bandwidth signal currents from a very low input impedance port to a high output impedance port. These devices are especially useful for the application under study: optical signal reception in FSO receiving systems. The proposed optical front-end design uses a CMD between the photo-detector and the transimpedance amplifier, in order to reduce the bandwidth limitation posed by the photo-detector's intrinsic capacitance acting together with the transimpedance amplifier's input impedance. Several guidelines for CMD

conception were advanced, based on system requirements following two general set-up configurations: TIA+CMD and Global Feedback TIA+CMD designs. The text then discussed CMD circuit design strategies based on CC-II, design and frequency optimization of class AB CC-II suitable for CMD operation, and methods to further reduce input impedance on CC-IIs. These final sections used a normalized CC-II model developed from circuit nodal analysis results. This normalized model allowed general circuit design considerations to be taken, irrespectively to the technological frame under choice. The final conclusion is that CMDs using a positive feedback CC-II configuration should be used in general for these applications.

6 EXPERIMENTAL VALIDATION

“You’re walking and you don’t always realize it but you’re always falling. With each step, you fall slightly forward and then catch yourself from falling. Over and over, you’re falling and then catching yourself from falling. And this is how you can be walking and falling at the same time.” – Laurie Anderson.

Summary

This chapter presents experimental and simulation results supporting the theoretical work described on the previous chapters.

The concept of delayed feedback was validated using two different approaches: integrated circuit simulation and experimental verification. The experimental validation was especially focused on the observation of the dynamics of a real delayed feedback amplifier and the accuracy of the theoretical perspectives discussed on chapter 3 and 4.

Integrated circuit simulations were made in order to validate the usage of both delayed feedback and active current matching, as suitable techniques for FSO transimpedance front end design. A simple transimpedance amplifier was adopted as reference for comparison against the optimized versions using both techniques.

6.1 Introduction

This chapter describes the experimental and simulation results achieved using the previously described bandwidth enhancement techniques.

The concept of delayed feedback was demonstrated using two different approaches. The first approach consisted on an experimental validation of the concept, using standard discrete electronic devices, since this technique is general enough to be applied to discrete circuits. This experiment was especially conceived to demonstrate and verify the concept theoretical perspectives described on chapters 3 and 4. Additionally, it served to demonstrate that common high frequency design guidelines demanding small and compact circuits are not always the best design approach. When concerning delayed feedback amplifiers it is wise to think in long feedback paths rather than small paths. From the bandwidth point of view this is a wise decision.

Since this experimental validation is not a good example of the potential of the delayed feedback concept, another evaluation using integrated circuit design was also devised. This evaluation validates the concept of delayed feedback as a useful bandwidth enhancement technique suitable for integrated circuit design. For this purpose, a generic transimpedance amplifier was designed and optimized using the available SiGe HBT/CMOS 350nm integrated circuit process from Austria Micro Systems. In order to cope with the initial motivation of this work, this transimpedance amplifier was designed using as reference common FSO receiver requirements, namely: large gain capabilities (larger than $50dB\Omega$), large bandwidth (larger than $10MHz$) and large intrinsic capacitance photo-detectors (between $1pF$ and $150pF$). These values were used to produce a reference amplifier, suitable to improvement by both delayed feedback and active current matching. The achieved results are a clear evidence of the validity of the proposed techniques.

Finally, the concept of active current matching was demonstrated using a simulation approach. In this respect, the same reference transimpedance amplifier used for delayed feedback demonstration was also used in this context. The achieved results produced a comparative basis between the two methods.

6.2 Delayed Feedback - Experimental Validation

This section describes the experimental set-up and the results achieved with the experimental validation of the delayed feedback concept. Due to its inherent simplicity, the delayed feedback concept can be easily demonstrated using standard discrete electronic components. In order to have a clear view of all the dynamic intricacies of the method, it was selected a well behaved voltage feedback configuration based on the traditional TL082 OA, employing a matched 50Ω coaxial cable

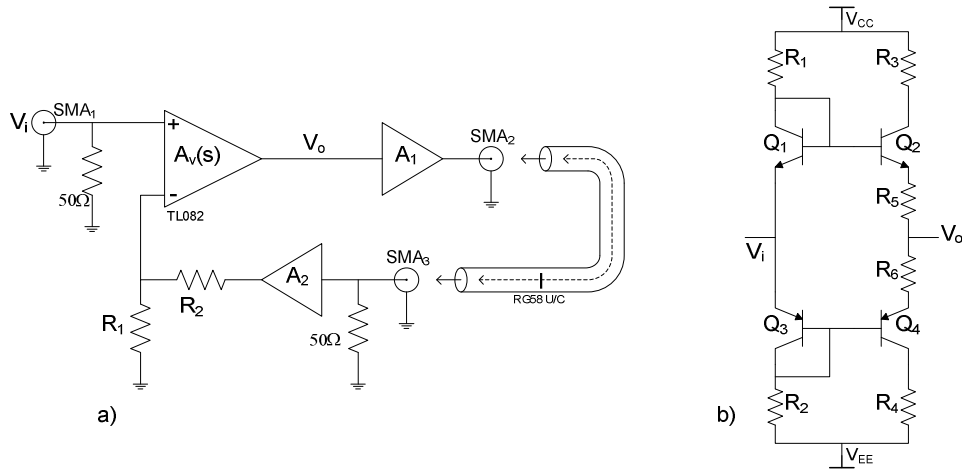


Figure 6.1 – Delayed feedback amplifier: a) experimental set-up and b) buffer circuitry.

as delay element.

6.2.1 Experimental Set-Up

Figure 6.1 a) depicts the experimental set-up devised to demonstrate the delayed feedback concept. The feedback amplifier consists of a voltage feedback configuration based on a TL082³⁹ OA and resistances R_1 and R_2 . Ignoring the delay element, composed by two voltage buffers (A_1 and A_2) and the transmission line (TL), this is a traditional non-inverting amplifier configuration. In order to synthesize several values of the return ratio, thus simulating the loop-gain of the ideal delayed feedback concept, resistance R_1 consists of a potentiometer and R_2 has a fixed value of 510Ω .

The delay element consists of a variable length TL, made of (3 to) 36 meters length RG58 U/C 50Ω coaxial cable. According to the discussion on delay element design taken in chapter 4, it is absolutely necessary to have a perfect match at the load end of the line in order to produce a perfect delay element. Buffer A_2 is used for this purpose; its high impedance in parallel with the 50Ω termination resistance, provides the necessary load matching. It is not required to have the same matching condition on the source end of the line. The source termination produces two distinct effects: power division between the source and line impedances; and delay distortion (when the load is not matched). The only effect to take into account in this case is power division, which can be minimized using a low output impedance buffer. As the TL082 output impedance is not small enough (typically around 50Ω), buffer A_1 provides the required small source impedance condition, thus reducing the power division at the source end of the line.

³⁹ The TL082 was chosen instead of the TL081 due to its robust built-in voltage offset control. The TL081 requires external circuitry to adjust voltage offset.

Figure 6.1 b) depicts the high bandwidth buffers used. These buffers are implemented using two class AB output amplifiers comprising both NPN and PNP BJT transistors 2N2222A (Q_1 and Q_2) and 2N2907A (Q_3 and Q_4), respectively. Resistances R_1 and R_2 together with the power sources V_{CC} (10V) and V_{EE} (-10V) set the bias condition of the buffer; for the present case these resistances had the value of $1k\Omega$, thus assuring a biasing current on the input branch (Q_1 and Q_2) of $9.3mA$. Resistances R_3 and R_4 act as short-circuit protection (both 100Ω). Finally resistances R_5 and R_6 assure class AB operation (both 10Ω). Under these conditions, these buffers had a voltage gain of $A_v \approx 1$, over a bandwidth of $BW \approx 40MHz$ (which is more than sufficient for the system under analysis). The low frequency input and output impedances exhibited a resistance value of $R_i = 500\Omega$ and $R_o = 10\Omega$, respectively. These values were obtained using PSPICE simulation and were also experimentally checked.

The experimental evaluation used a HP4195A network analyzer and a HP41800A 500MHz high impedance active probe. The calibration was done using a thru connection between input and output, and a resolution bandwidth (RBW) of $10Hz$. Most of the measures were done using less demanding resolutions in order to accelerate the measuring process. Nevertheless, the initial RBW of $10Hz$ was used for situations demanding higher resolutions, as is the case of oscillating conditions. The next section presents and discusses the measured results.

6.2.2 Numerical Model

The experimental procedure consisted on the measurement of BWER for different return ratio and TL lengths. A reference set-up consisting on the amplifier of Figure 6.1 a) with a residual TL of $50cm$ was used in order to provide a cut-off frequency reference value. The BWER was then computed against this reference value. The measured BWER values were then compared with the theoretical model obtained using Matlab simulation. The theoretical model followed the design analysis guidelines discussed on chapter 4. Equation (4.36) was used to describe the delay model of the terminated TL. Since the load termination is matched to the TL characteristic impedance, it follows that $\rho_L = 0$, thus simplifying the delay model to,

$$H_D(s) = \frac{Z_o}{Z_o + Z_s} e^{-\gamma(s)l} \quad (6.1)$$

where Z_s represents the source impedance (for this case Z_s is the buffer's output impedance, assumed frequency independent for the frequencies of interest). $\gamma(s)$ and l represent the propagation constant and the length of the TL, respectively. The propagation constant, $\gamma(s)$, was obtained from the characteristics of the RG58 U/C coaxial cable, namely: characteristic impedance, $Z_o = 50\Omega$,

relative permittivity, $\epsilon_r=2.3$; negligible loss below 10MHz (less than 0.044dB/m). Taking this into consideration, the phase constant becomes,

$$\gamma(j\omega) = j\omega \frac{l\sqrt{\epsilon_r}}{c} \quad (6.2)$$

where c represents the velocity of light.

The feedback amplifier model was obtained using Rosenstark's method (section 4.3.2) and the return ratio definition. Defining the gain of the TL082 OA as the control parameter and the output voltage V_o as the output parameter, this leads to

$$\begin{aligned} R(s) &= \frac{R_1}{R_1 + R_2} A_o(s) A_{v_1} H_D(s) A_{v_2} \\ G_\infty(s) &= \left(\frac{R_1}{R_1 + R_2} A_{v_1} H_D(s) A_{v_2} \right)^{-1} \\ G_0 &= 0 \end{aligned} \quad (6.3)$$

where $A_o(s)$ is the gain of the TL082 OA, A_{v_1} and A_{v_2} are respectively the gains of buffers A_1 and A_2 . Finally, the closed-loop gain becomes,

$$A_{v_f}(s) = G_\infty(s) \frac{R(s)}{1 + R(s)} = \frac{A_o(s)}{1 + R(s)} \quad (6.4)$$

Equations (6.1) to (6.4) describe the model used to predict the BWER values in Matlab. All the real experiment conditions were recreated using this approach. According to chapter 3 the loop-gain is the quantity of interest. Using this modeling approach, the loop-gain is represented by $R(0)$, which can be experimentally evaluated using $A_o(0)$ and $A_{v_f}(0)$. It was assumed (and experimentally confirmed) that the TL082 has an open-loop voltage gain of $200k$ and two relevant poles, at the frequencies of 10.5Hz and 6MHz (values obtained experimentally, using the reference set-up).

6.2.3 Results and Discussion

Figure 6.2 represents the BWER as a function of the return ratio, for various lengths of the TL (various delays), using as reference the amplifier with a small 50cm long TL. The solid lines in Figure 6.2 represent values for a second order TL082 model obtained using numerical simulation in Matlab. The dots represent the experimental results for various TL lengths. The matching between theoretical and experimental results is evident. Minor deviations from the theoretical model are due to insufficient detail on both OA and TL models. Figure 6.3 shows that the maximum BWER under the maximal flat frequency response constraint is 2 (corresponding to 100% increase). Nevertheless, the BWER increases further above this condition. According to Figure 6.2 the theoretical maximum is slightly above 2.2, which agrees with the experimentally observed maximum of 2.19, observed

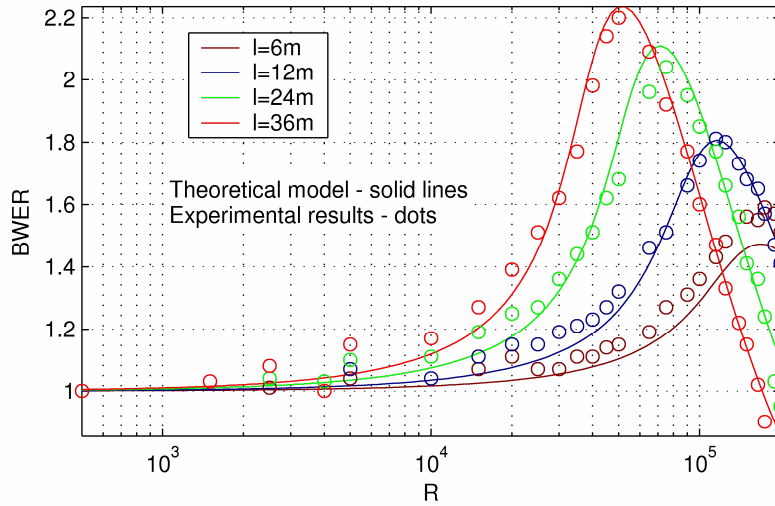


Figure 6.2 – BWER experimental results.

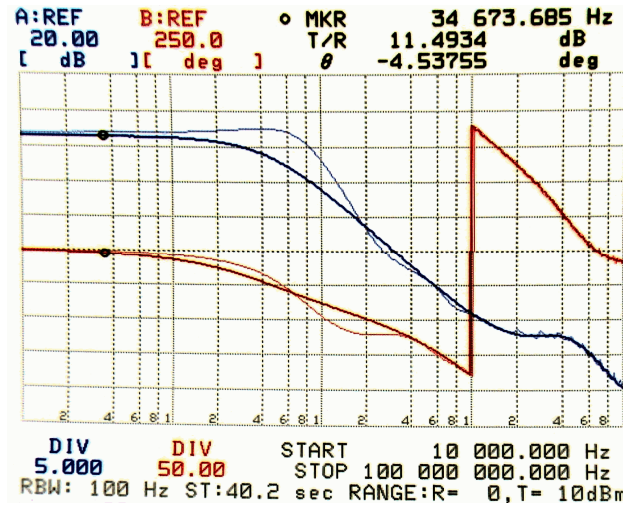


Figure 6.3 – Frequency response: maximum BWER observed under maximum flat condition.

with a TL line measuring $36m$.

Frequency peaking was observed for large values of the return ratio and all the TL lengths, as predicted in chapter 3. Figure 6.4 is an experimental evidence of this fact. Figure 6.4 also shows the bounded nature of the frequency response, when $R(0)$ is fixed and the delay increases (increasing the length of the TL, l). BWER increases while there are small peaking effects. Further increases on l (or alternatively, on $R(0)$) beyond this condition turn the amplifier less stable and also decrease BWER. In fact the frequency response corresponding to the descending part of the curves on Figure 6.2 exhibited large values of frequency peaking.

The oscillation condition was even observed for the maximum length TL tested ($36m$), as depicted on Figure 6.5. This figure shows that the frequency response under oscillation exhibits a multiple maxima behavior, consistent with the theoretical results of chapter 3. Nevertheless, it is interestingly observed that for this condition the system no longer obeys to the bounded frequency lemma (this result from the comparison of Figure 6.4 and Figure 6.5). This can be explained

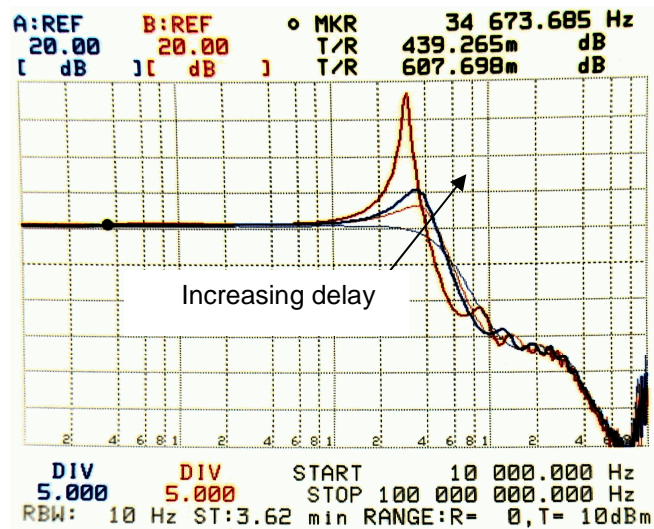


Figure 6.4 – Bounded frequency response nature.

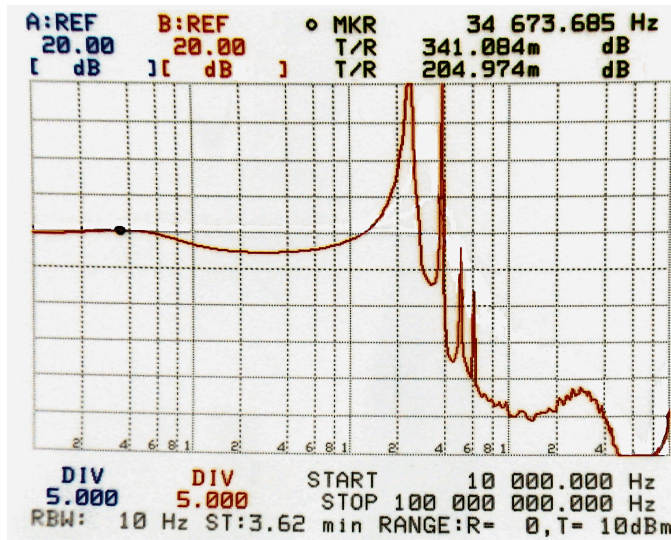


Figure 6.5 – Magnitude of the frequency response: oscillation condition.

recalling that under oscillating conditions the small signal linear model is no longer valid, thus preventing the bounded frequency response requirements to be fulfilled.

This simple example clearly shows all the important conclusions mentioned on chapters 3 and 4, namely:

- The BWER is strongly dependent on the length of TL used in the experiment, revealing the strong dependence on the associated delay;
- The magnitude of frequency response of a delayed feedback amplifier is indeed (usually) bounded.
- Oscillation due to the excess phase introduced by the delay element was verified even with a well characterized, almost one pole amplifier, thus proving that even a first order feedback amplifier can become unstable due to delay terms inside the feedback loop;
- Finally, when concerning feedback amplifiers it may be wise to think in long feedback paths

rather than small paths. From the bandwidth point of view this is a wise decision. For this example it was observed larger bandwidths for the amplifier with the $36m$ long TL than for the $50cm$ long case!

This practical demonstration should be seen as experimental confirmation of the theoretical results presented on chapters 3 and 4 and not as an exploitation of the concept. The next sections explore the usage of this concept to design frequency optimized transimpedance amplifiers using integrated circuit technology. A subset of these results was previously presented in two international conference papers [296, 297]. A refinement of these contributions was published in the International Journal of Circuit Theory (currently available as an early view electronic document, waiting to be printed) [301].

6.3 Delayed Feedback – Integrated Circuit Implementation

This section presents simulation results of the design and bandwidth optimization procedures of a transimpedance amplifier. This circuit was design using the 350nm SiGe HBT/CMOS process from Austria Microsystems. The main objective of this design was to demonstrate the usage of the concept of delayed feedback in integrated circuit design. This evaluation recurred to schematic level simulation, allowing the analysis of the complete dynamics of the amplifier for different design conditions.

It was also important to cope with the overall motivation of this work, the design of high-gain and high-bandwidth transimpedance amplifiers suited for FSO receiving systems. As such the design used a transimpedance amplifier with fixed gain of $80dB\Omega$ and bandwidth optimized for different input capacitance conditions (ranging from $1pF$ to $150pF$).

6.3.1 Reference Transimpedance Amplifier Specifications

Figure 6.6 depicts the conceptual design of the transimpedance amplifier (TIA). The main amplifier consists on a three stage amplifier: a differential input stage, an inverting second stage (also used as delay element) and an output stage. Figure 6.6 also represents the small signal equivalent model of the photo-detector, consisting in a current source I_P in parallel with a

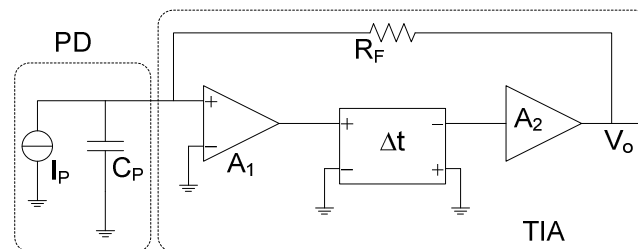


Figure 6.6 - TIA conceptual design.

capacitance C_P .

Transimpedance amplifiers based on single transistor input stages present some biasing difficulties. Since the biasing is preferably established through the feedback resistance, this results in an added restriction on the possible values of R_F (and hence, closed-loop gain). A possible solution to reduce this problem is to use differential input stages, and fixing the bias level in one of its inputs.

The main amplifier must include a delay element in order to apply delayed feedback concepts. For this reason, the second stage is a delay element based on a CX configuration (as in section 4.4.3). As CX configurations have negative voltage gain, this implies one of two design choices:

- i) The TIA input is the positive input of the differential stage and A_2 is a non-inverting output buffer;
- ii) The TIA input is the negative input of the differential stage and A_2 is an inverting amplifier.

It is important to have a good output buffer in order to reduce feedback loading effects at the output port, thus the first configuration was selected.

In order to evaluate the delayed feedback effects in this amplifier it is necessary to have a reference amplifier, and to measure the BWER against this reference as the delay is adjusted. For that purpose, a reference TIA consists on the circuit of Figure 6.6 having its delay element set to its minimum. In this situation, it is possible to say that the reference TIA has only internal delay contributions. The reference amplifier was conceived in order to reflect traditional FSO systems design constraints, namely:

- i) Able to handle photo-detector with intrinsic capacitance ranging from $1pF$ to $150pF$;
- ii) Closed-loop gain adjustment capabilities, ranging from 500Ω to $10k\Omega$;
- iii) Bandwidth larger than $5MHz$ (for all C_P values);
- iv) A second pole placed above $200MHz$.

The first requirement assures the dominant character of the input pole. Gain adjustment capabilities were necessary to demonstrate delayed feedback effects, allowing the amplifier dynamics to change accordingly to the required gain (feedback resistance). A second pole placed above $200MHz$ provides large damping coefficients and consequently a clear observation of the bandwidth enhancement effects.

6.3.2 Overall Circuit Design

Figure 6.7 shows the transimpedance amplifier circuit used for this study. The main amplifier is composed by three amplifying stages connected on an overall transimpedance feedback configuration. The first stage uses an unbalanced differential pair, formed by two HBT transistors

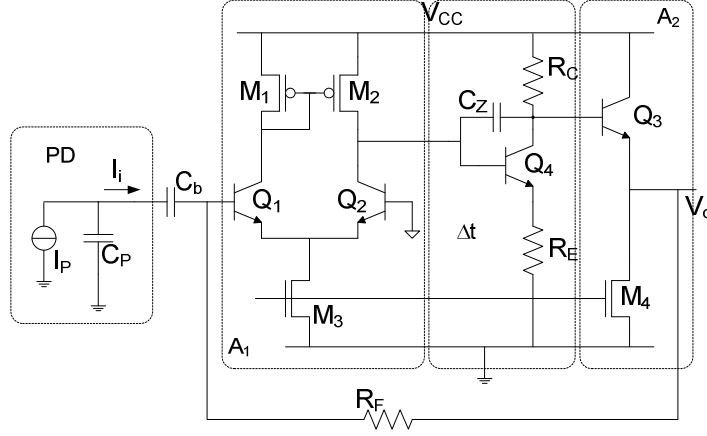


Figure 6.7 – Transimpedance amplifier circuit.

Q_1 and Q_2 , an active load made off PMOS transistors M_1 and M_2 , and a current biasing source M_3 . Transistor Q_1 has small voltage gain due to M_1 . This condition is especially important since it reduces the Miller effect on the input transistor. Q_2 acts as a common base amplifier, providing the necessary voltage amplification. Transistors Q_1 and Q_2 provide also a robust biasing scheme. Traditional single input transistor configurations have several biasing problems, thus preventing the usage of wide range feedback resistances. The voltage at the base of Q_2 provides the reference bias level, which must be followed by the amplifier's output at Q_3 .

The second stage is a simple degenerated common emitter stage. According to chapter 4, transistor Q_4 with its collector and emitter resistances form an all-pass delay element. In order to synthesize an approximated delay transfer function, resistance R_C and R_E must be closely matched. This condition also implies a voltage gain of -1 . Finally, transistor Q_3 forms the output stage. A common collector configuration was selected for this purpose in order to provide both, isolation means for the delay element and small output impedance.

The final circuit was biased with a 3.3V power source and exhibited a power consumption of 10.5mW. The power penalty due to the addition of the delay stage (transistor Q_4) was only 2.3%. Table 6.1 summarizes all the transistor specifications and biasing conditions. Figure 6.8 shows the final layout of the transimpedance amplifier. The presence of C_Z on the layout indicates that this is not the reference condition (external delay set to zero). In fact, this picture corresponds to an optimized version of the reference TIA, when the photo-detector intrinsic capacitance is 150pF. For this situation, the circuit occupied an area of $4.8nm^2$. The inclusion of a delay element represents an area penalty of 42% for this case. However, the area penalty is smaller for other values of C_P .

6.3.3 Simulated Delay Characterization

The first set of simulations was dedicated to the verification of the delay design approach presented on chapter 4. As presented (section 4.4.3), the ideal behavior of a complex exponential

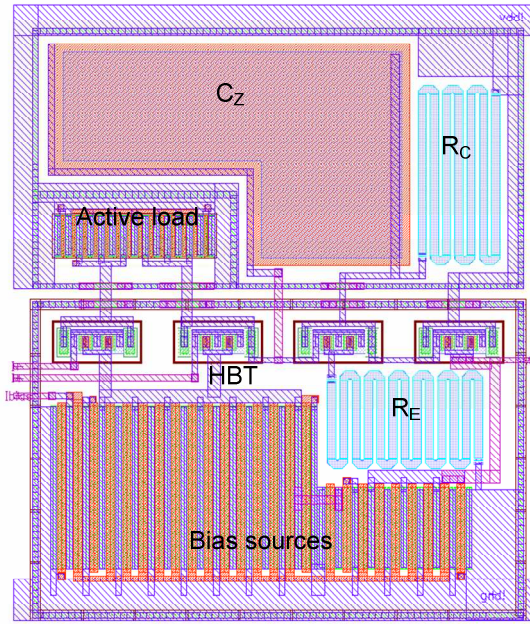


Figure 6.8 – TIA layout.

Table 6.1 – Transistor specifications and quiescent conditions.

	Type	I_{CQ}	V_{CEQ}
Q ₁	HBT npn232 $A_c=1.6$	1.2mA	989mV
Q ₂		1.2mA	813mV
Q ₃		550μA	1.824V
Q ₃		75μA	1.618V
		$ I_{DQ} $	$ V_{DSQ} $
M ₁	PMOS 30/0.35	1.2mA	1.608V
M ₂			1.784V
M ₃	NMOS 250/1	2.4mA	703mV
M ₄			1.682V

can be simulated accurately by a first order Padé approximant. These approximated delays were named all-pass delays, thus revealing their all-pass characteristics. It was also described how these all-pass delay elements could be designed using simple degenerated common X stages. The following simulation results provide a verification of the proposed modeling approach described on chapter 4. These results were previously presented on an international conference paper [299].

The simulation set-up consisted in the circuit of Figure 6.9. The generic transistor was substituted by an HBT, resistances R_C and R_E were fixed to meet both, unitary gain and optimum f_T .

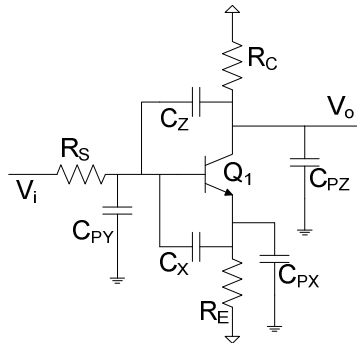


Figure 6.9 – Simulated delay element circuit.

The transistor bias point was set to $I_C=1.16mA$ and $V_{CE}=1.7V$, using a bias source $V_{CC}=3.3V$. The resultant simplified small-signal model had $g_m=40.9mS$, $\beta=149$, $C_\pi=162.1fF$ (C_X), $C_\mu=16.2fF$ (C_Z intrinsic part) and $\tau_i=120ps$. Several simulation tests were made, reflecting the influence of C_{PX} , C_{PY} and C_{PZ} on the phase and amplitude errors variation, when C_Z was varied (thus varying the effective delay).

The results are depicted on Figure 6.10, Figure 6.11 and Figure 6.12. For each case, it was assumed a fixed parasitic capacitance (C_{PX} , C_{PY} or C_{PZ}) of $100fF$, which represents an unusually large parasitic capacitance in a 350nm technology, but served better the purpose of model verification. Amplitude and phase errors were evaluated using the definitions in section 4.4.3.1 (equations (4.47) and (4.48)). It is noticeable that the phase error variation for these three cases

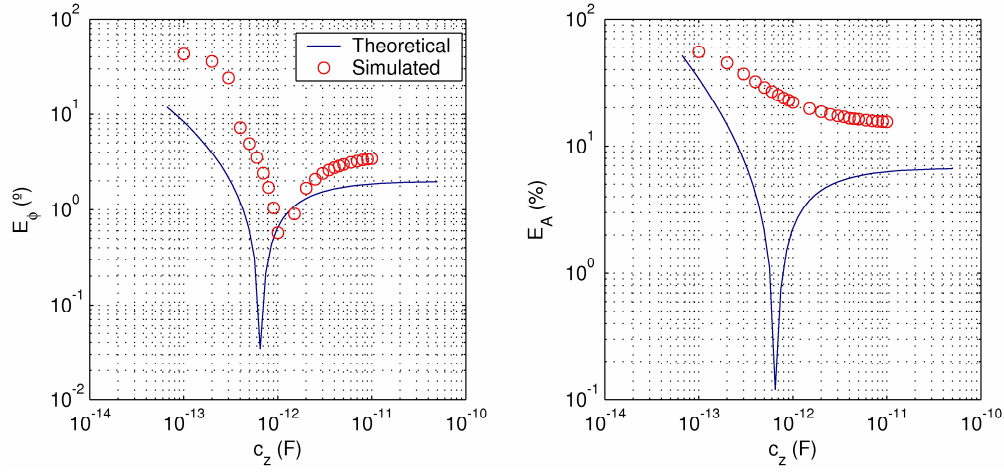


Figure 6.10 - Phase and amplitude error with fixed $C_{PX}=100fF$.

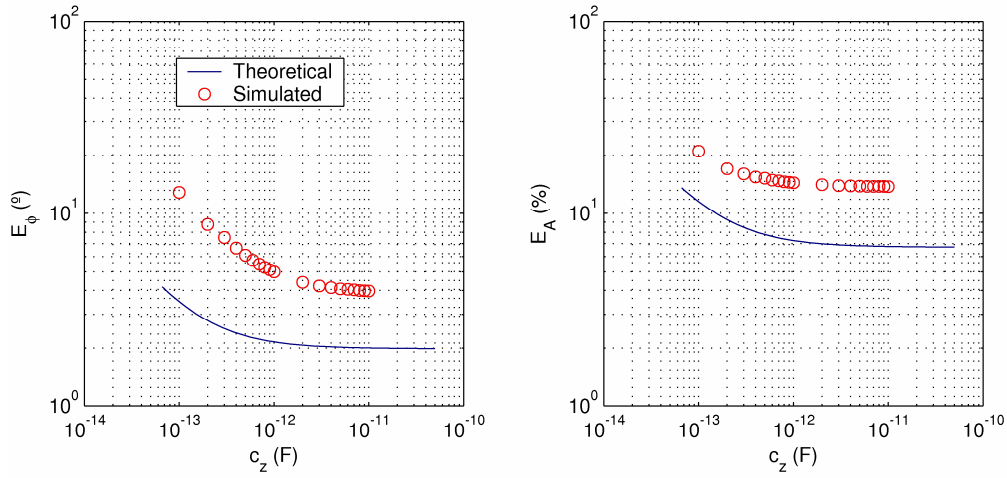


Figure 6.11 - Phase and amplitude error with fixed $C_{PY}=100fF$.

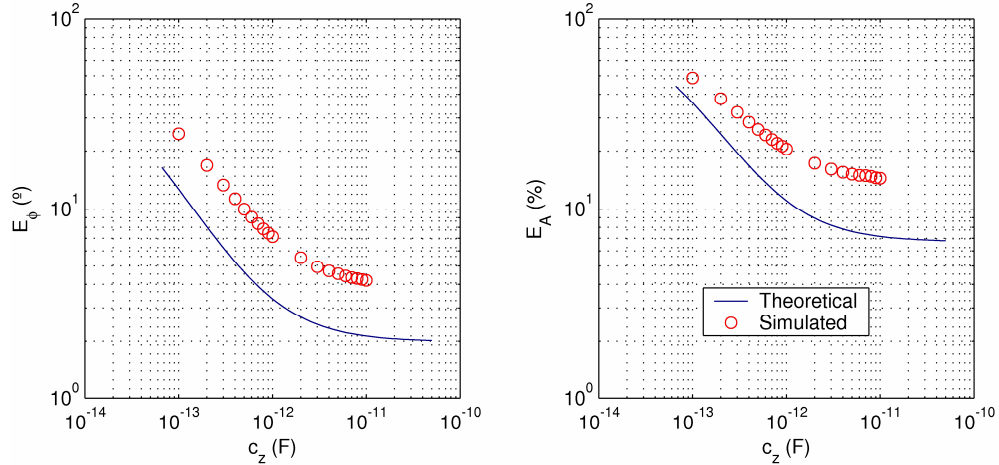


Figure 6.12 - Phase and amplitude error with fixed $C_{PZ}=100fF$.

follows closely the behavior predicted by the proposed model. For large values of C_Z the phase error tends to a constant value. This effect was observed with a maximum deviation of 2° , meaning that the phase characteristic is indeed in agreement with the required conditions to meet the desired delay. Small values of C_Z (near the parasitic values assumed) lead to large phase errors and higher

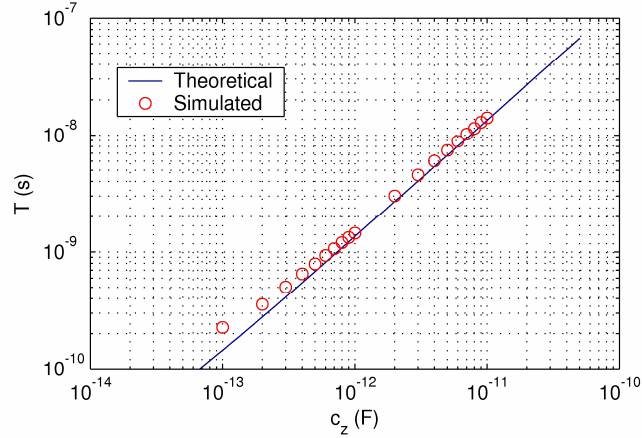


Figure 6.13 - Measured and computed time delay.

deviations from the model. This is due to the higher order dynamics of this circuit. It was also verified the presence of a minimum phase error behavior for the case of fixed C_{PX} . The amplitude error was always larger than the model prediction. This is also due to the presence of higher order terms in the circuit transfer function. In general the amplitude response exhibited the predicted step like response. However, the presence of more poles and zeros also implies that for increasing frequencies the amplitude rolls off. This makes the procedure for measuring the amplitude error discussed on section 4.4.3.1 prone to imperfections. Nevertheless, this simplified model allows good predictions of the effective time-delay, as depicted on Figure 6.13. The maximum error was of 20%, verified only for small values of C_Z , as expected. This could be explained simply by noting that for small values of C_Z , the effect of parasitic capacitances becomes noticeable, especially on the phase error. Since, the phase error has larger impact on time-delay prediction, it is expectable to experiment larger errors for these situations.

6.3.4 Delayed Feedback Verification

The usage of delayed feedback in integrated circuit amplifier design poses several design challenges. First of all, using long transmission lines as delay elements is not a viable strategy (except perhaps, for the cases when either the return ratio or the open-loop poles have large values). Secondly, it is usually difficult to achieve large return ratios independent of the feedback configuration in use. Thirdly, maximizing bandwidth usually compromises the return ratio.

The second and third problems are especially relevant in transimpedance amplifiers, as discussed on chapter 4. When a high capacitive source is feeding the transimpedance amplifier, both return ratio and dominant pole (arising due to the input circuitry) are strongly dependent on the feedback resistance. According to the results of chapter 4 a suitable model to evaluate the effect of the feedback resistance on the return ratio and open-loop poles must consider the amplifier open-loop

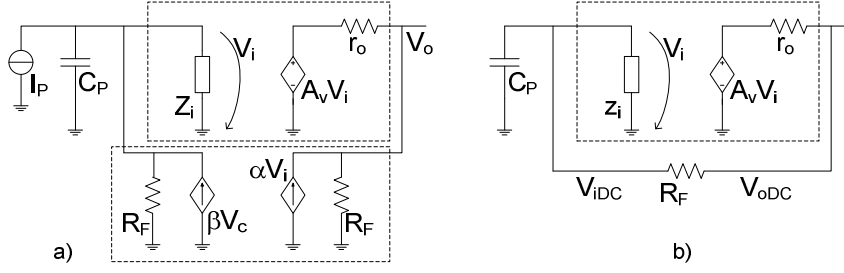


Figure 6.14 – Transimpedance amplifier: a) external source model representation; b) bias model.

voltage amplification, the input impedance (or equivalently, the open-loop transimpedance, including both voltage amplification and input impedance) and possibly, the output impedance (when this takes values comparable to the feedback resistance). The following results were compared with the theoretical model predictions, based on these considerations.

The first task was to measure the amplifier dynamics, observing both open-loop and closed-loop conditions. Measuring the open-loop dynamics of a feedback amplifier is not a simple task. In general the feedback network affects both small signal and large signal operation of the internal amplifier. Transimpedance amplifiers are an example of such non-ideal behavior. The reason for this difficulty lies on the fact that the input stage biasing is done through the feedback resistance. Thus, opening the feedback loop cannot be accomplished using the simple feedback loading concept of the traditional β -A approach. Using the general return ratio definition is also inappropriate for the present case, since none of the transistors (Q_1 to Q_4) provides a good controlled reference source (it would be also difficult to break the loop on an internal transistor without compromise biasing).

An efficient way to overcome these difficulties while coping with the general return ratio analysis consists on the method reported in [105]. According to this method, the feedback network can be converted into an internal or external source model (see appendix C 1 for a short description). For the present case, the external source model was adopted resulting in the set-up of Figure 6.14a, where α and β are loop control parameters with reference value equal to $1/R_F$.

Breaking the feedback is now a simple matter. It suffices to use a test set-up consisting of the two circuits depicted in Figure 6.14: the equivalent external source circuit (Figure 6.14a) for small signal analysis; and an auxiliary operating point equivalent circuit (Figure 6.14b), consisting on the original amplifier without any independent signal source. Selecting β as the controlled parameter, closed-loop operation is selected when V_c is set to V_o , on the other hand, open-loop operation is selected using $V_c=V_{oDC}$. This procedure assures that, when $V_c=V_{oDC}$ the amplifier operates under closed-loop conditions from a biasing perspective, but it is at the same time under open-loop condition for signal concerns.

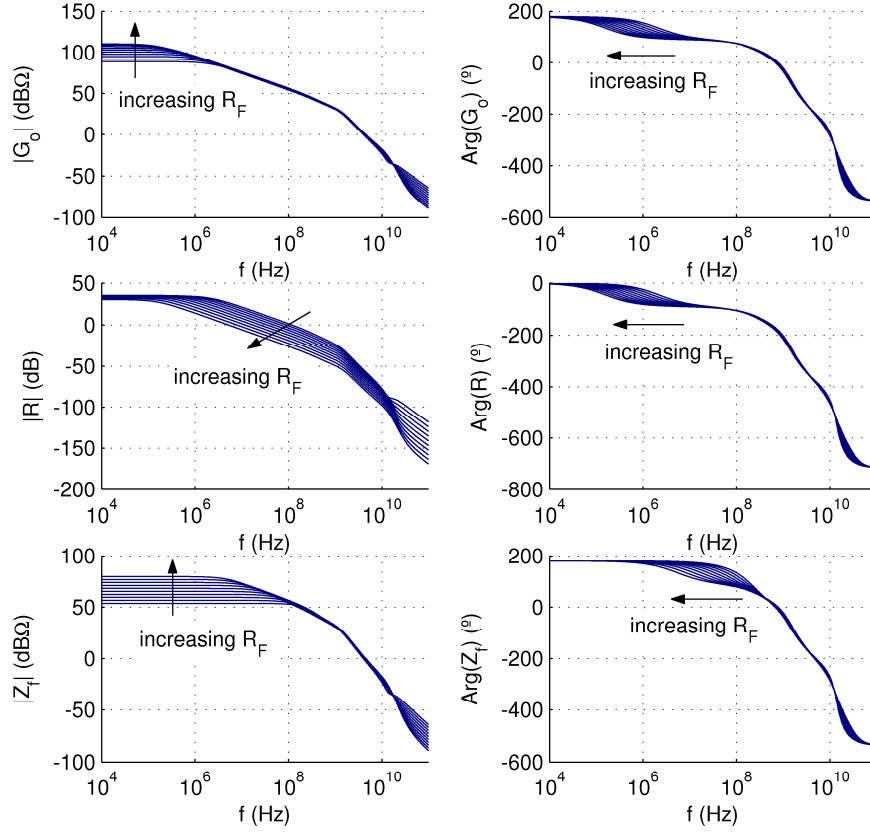


Figure 6.15 – TIA small signal measurements: from top to bottom, direct gain, return ratio and closed-loop gain (magnitude and phase).

Selecting β as the control parameter, V_o as both output and control variable and considering the amplifier simplified model of Figure 6.14 (consisting on voltage gain $A_v(s)$, input impedance $z_i(s)$ and output resistance r_o), the return ratio is given by:

$$R(s) = \left(A_v(s) - \frac{r_o}{R_F} \right) \frac{R_F}{r_o + R_F} \frac{z_i(s)}{z_i(s) + R_F} \quad (6.5)$$

Following the usual definitions, $G_o(s) = -R_F R(s)$ and $G_\infty(s) = 0$, the closed-loop gain becomes

$$Z_f(s) = -R_F \frac{R(s)}{1 + R(s)} \quad (6.6)$$

The term $A_v(s)$ is usually larger than r_o/R_F , especially for large values of R_F . It is also clear that $A_v(s)$ includes all the terms contributing for the voltage gain between V_i and V_o , thus it necessary includes a delay transfer function as given by equation (4.54). Figure 6.15 depicts the simulation results of both open-loop return ratio and direct gain (selecting β as the controlled parameter) and closed-loop transimpedance gain measurements, for various values of R_F . It is readily apparent that $G_o(s)$ and $R(s)$ differ only from a constant factor, both transfer functions have exactly the same pole contributions (this is apparent from the phase plots of $R(s)$ and $G_o(s)$). These results show that closed-loop operation is stable for the assumed values of R_F , ranging from 500Ω to $10k\Omega$.

The theoretical model for delayed feedback prediction was based on these open-loop simulations. This model must accurately predict all the open-loop dynamics. Assuming that $z_i(s)$ is accurately expressed by a capacitance c_i in parallel with a resistance r_i , the open-loop input pole is given by,

$$p_1 = \frac{r_i + R_F}{r_i R_F (c_i + C_P)} \quad (6.7)$$

It is noticeable that c_i adds in parallel to C_P to establish the input pole. In this example it is assumed that C_P represents the dominant part, while c_i is the input capacitance of Q_I (very small when compared to C_P). The input pole depends also on the parallel association of r_i and R_F , the open-loop input impedance with feedback loading. Looking again to (6.5) it is readily seen that both p_1 and $R(0)$ are depend on R_F . Thus, in order to apply the results of chapter 3 and 4 to predict the BWER it is necessary to express both p_1 and $R(0)$ as functions of R_F .

The return ratio and p_1 were based on equations (6.5) and (6.7), relying on measured perspectives of $A_v(0)$, r_i and r_o . Table 6.2 depicts the measured values of $A_v(0)$, r_i and r_o . As can be seen, these values exhibited small variations for different values of R_F . These variations occur because changing R_F also affects the biasing of Q_I .

Table 6.2 – Open-loop measured dynamics.

	Min	Avg	Max	units
$A_v(0)$	48.9	57.6	63.1	–
r_i	14.5	16.7	20.3	$k\Omega$
r_o	–	65	–	Ω
p_2	283	304	332	MHz

It was observed that both p_2 (the second pole) and r_i increased with increasing values of R_F . The same monotonic behavior was not observed for $A_v(0)$ which exhibited a local maximum (63.1) for R_F values between 500Ω and $10k\Omega$. Using the average values of r_i , $A_v(0)$, R_F and C_P it was confirmed that equations (6.5) and (6.7) fitted the values of $R(0)$ and p_1 with an absolute error of less than 10%.

Figure 6.16 shows the comparison of the theoretic perspectives of BWER and $R(0)$ against the simulated results. The BWER was measured using a two step procedure. First, the bandwidth of the reference amplifier having $C_Z=0$ was measured for several values of R_F . Then, the bandwidth was optimized varying C_Z in order to reach the maximum flat condition (frequency overshoot of less than 0.5% was tolerated), for the same set of R_F values. The values of C_Z were registered and used to determine the loop delay. It is noticeable the proximity of the theoretical perspectives even with insufficient detail of the amplifier dynamics. As the BWER plot of Figure 6.16 shows, the simulated results are accurately fitted between the minimum and maximum models, using minimum and maximum perspectives of $A_v(0)$, r_i and p_2 , respectively. This is consistent with the fact that all these quantities vary for different values of δ . It is also noticeable that the majority of the measured results are accurately predicted by the average model. The return ratio dependence on $A_v(0)$ is

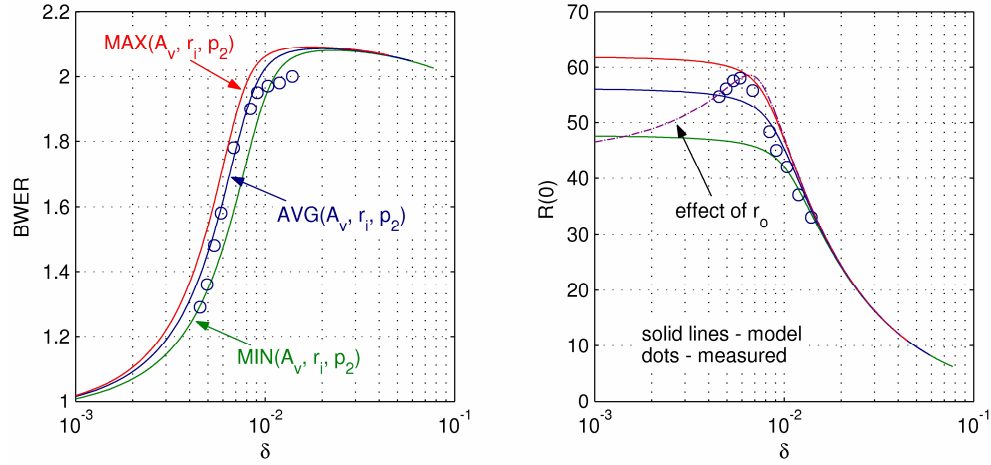


Figure 6.16 – BWER and return ratio theoretical results versus simulated results.

patent on the non-monotonic distribution of the measured results. This behavior was not observed for any of the three modeling approaches, using minimum, average and maximum values of $A_v(0)$, r_i and p_2 .

According to (6.5) it is possible to use r_o to model the variation of $A_v(0)$ for different values of R_F . In fact using larger values of r_o , increases voltage division on the output port due to R_F . This effect is similar to the observed gain variations, as depicted on Figure 6.16 ($R(0)$ plot). However, this contradicts simulation results, since r_o remain essentially fixed for all tested values of R_F .

6.3.5 Post Layout Simulation

The results reported on the previous section were obtained using schematic level simulations. This approach was there adequate, since the number of design variables under investigation (R_F , C_Z and C_P) was high enough to force many changes on design prior to each simulation step. However, since the layout parasitic elements can in general affect the amplifier's performance, it is of the utmost importance to verify circuit's operation through post layout simulation.

Several simulations were carried out with an extracted version of the layout of Figure 6.8. These simulations served two purposes: i) observing the effect of parasitic capacitances on delayed feedback; and ii) inspecting the range of C_P for which this amplifier provides a viable solution.

The reference set-up consisted of the base amplifier employing a feedback resistance R_F of $10k\Omega$, thus providing both closed-loop and open-loop fixed conditions. Open-loop measurements reveal that $A_v(0)=50.43$, $r_i=20.4k\Omega$ and $p_2=204.2MHz$ (the second pole was severely affected by the parasitic capacitances). Measured values of $R(0)=33.84$, $G_o(0)=338.4k\Omega$ and $Z_f(0)=9.71k\Omega$ are in complete agreement with (6.5) and (6.6). In order to have some means of simulating variable pole delay products, the simulation set-up used several values of source capacitance C_P , ranging

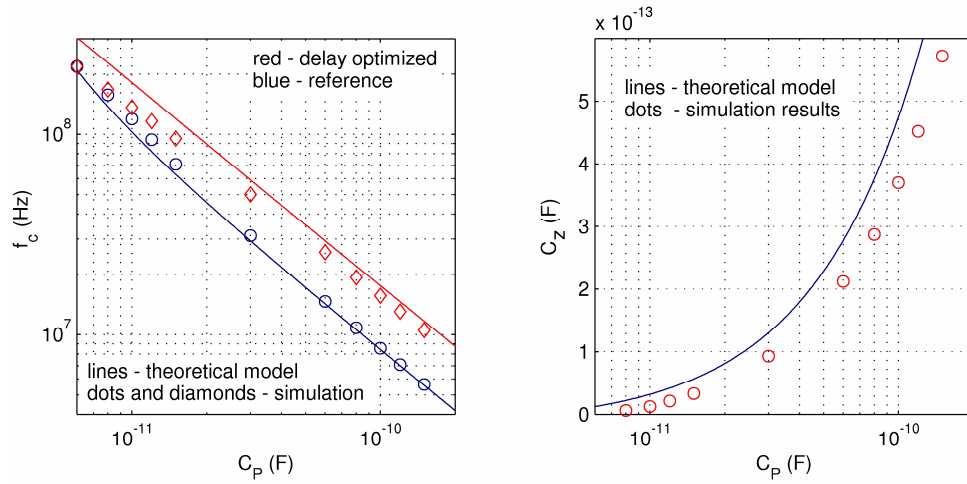


Figure 6.17 – Cut-off frequency and optimizing values of C_Z .

from $1pF$ to $150pF$. For this set of C_P values, having $C_Z=0$ (removing C_Z from the layout) the input pole p_1 data fit accurately the predictions of equation (6.7). The next simulation set consisted on an optimizing step: for each C_P , the value of C_Z was varied in order to achieve the maximum flat condition (as before, frequency overshoot of less than 0.5% was tolerated).

Figure 6.17 shows the results of both theoretical predictions and measured values of the cut-off frequency and the necessary optimizing C_Z , as functions of C_P . The theoretical model is more accurate for higher values of C_P . This can be explained by the fact that large values of C_P require larger delays to achieve maximum flatness, for which the effect of small parasitic capacitances introduce negligible errors. This is also evident for the predicted C_Z values: the measured C_Z values were always less than the theoretical predictions due to the influence of parasitic capacitances. The same effect was not so dramatic for the previous schematic simulations. Nevertheless, the theoretical model predicts accurately the cut-off and C_Z tendencies as displayed in Figure 6.17.

Figure 6.17 depicts results of the cut-off frequency and C_Z for values of C_P larger than $6pF$. For values of C_P of less than $6pF$ the frequency response revealed significant frequency peaking even without the optimizing with $C_Z=0$. This is also evident from the cut-off frequency plot, where the measured values of the reference amplifier ($C_Z=0$) meet the values of the optimized version. For these situations, the combined effect of the internal delay and pole frequencies impairs stability. Adding more delay to the loop for these situations would result in an unstable design.

Noise minimization was not taken into consideration, nor from a theoretical point of view (as part of the previous chapters) nor as a relevant design consideration for this proof of concept design. Nevertheless, due to its importance on the overall amplifier performance, some noise analyses were carried out in order to validate this design perspective. Figure 6.18 depicts the equivalent input noise spectral density frequency behavior for several values of C_P under the optimized maximum flat condition. As can be seen, the optimization procedure does not affect the noise floor, which

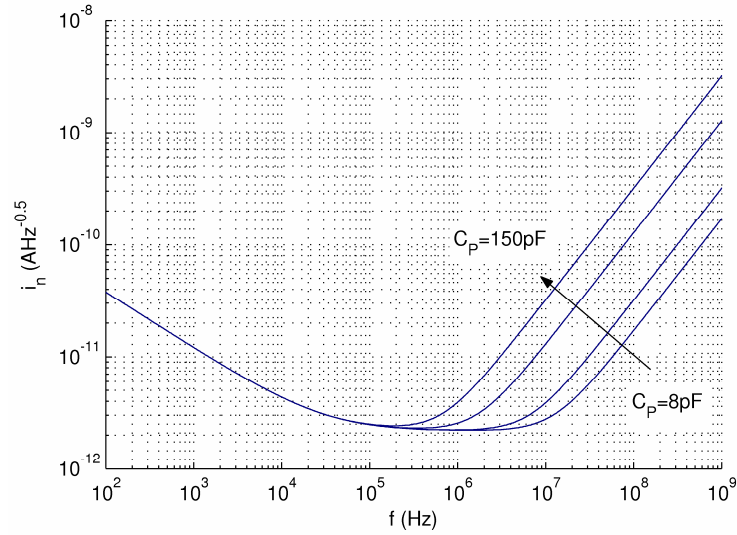


Figure 6.18 – Noise frequency response.

remained near $2.2pAHz^{-0.5}$. There is however a slight increase of the median noise current due to the fact that the circuit bandwidth was increased.

6.4 Active Current Matching Concept Validation

The previous sections showed that, delayed feedback is a valuable tool for bandwidth optimization in feedback amplifiers. This concept was then applied to the design and optimization of a transimpedance amplifier especially suited for FSO receiving systems. Following the discussion in chapter 5, active current matching can be employed as another means of achieving higher bandwidth in these transimpedance amplifiers. This section provides simulation evidence of the usage of active current matching as an effective bandwidth enhancement method.

6.4.1 Simulation Set-up

Figure 6.19 depicts the simulation set-up used to evaluate the active matching concept performance. This circuit comprises two main blocks: the current matching device (CMD) and the transimpedance amplifier (TIA). The CMD stage is represented as a current conveying element between the input port (connected to the photo-detector) and the output port (connected to the transimpedance stage). Chapter 5 explored this concept and presented several CMD design

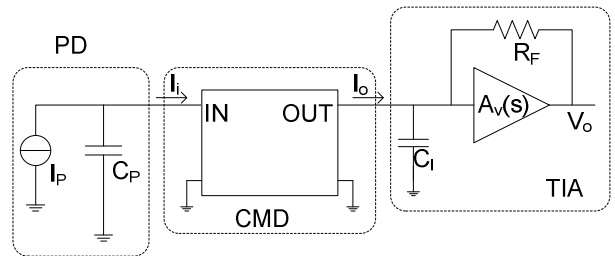


Figure 6.19 – TIA plus CMD simulation set-up.

approaches based on feedback configurations using CC-IIs. The transimpedance amplifier consisted on the reference configuration reported in section 6.3.2. The circuit of Figure 6.19 also shows the photo-detector's small signal equivalent. The TIA block depicts an interstage capacitance C_i ,

needed for compensation purposes, as discussed in section 6.3.5.

The reference design adopted for this study was the same as in section 6.3.1, that is, a reference amplifier suited to FSO receiving systems with the following capabilities:

- i) Able to handle photo-detector with intrinsic capacitance ranging from 1pF to 150pF ;
- ii) Closed-loop gain of $10\text{k}\Omega$;
- iii) Bandwidth larger than 5MHz (for all C_p values);
- iv) A second pole placed above 200MHz .

This section addresses the validation of the purposed active current matching technique. According to Figure 6.19 there are three design steps needed for this proof of concept:

- i) CC-II design and optimization. According to section 5.3.2 CC-II are suitable candidates for CMD design;
- ii) CMD optimization, based on the results of section 5.5 it is necessary to have a clear understanding of the impedance control using CC-II based CMDs;
- iii) Overall system performance evaluation, accomplished using post layout simulation.

The first and second tasks adopted a schematic simulation perspective due to the high number of design variables involved. Following the design guidelines presented on chapter 5, a class AB CC-II was chosen for this demonstration. CMD design and optimization was based on a CC-II employing positive feedback to reduce input impedance. Finally, the layout step addressed system design issues, providing a framework for performance evaluation of both, TIA (using the results of section 6.3.5 as reference) and TIA+CMD design perspectives.

The following sections are organized according to these three tasks. Section 6.4.2 addresses CC-II design issues. Section 6.4.3 addresses CMD performance evaluation issues. Finally, section 6.4.4 presents a comparative study between delayed feedback optimized TIA, and TIA+CMD design approaches. Section 6.4.5 presents some concluding remarks concerning the usage of these techniques.

6.4.2 CCII Circuit Design and Optimization

Figure 6.20 depicts the schematic of the class AB CC-II chosen for this evaluation framework. This circuit is a traditional translinear loop CC-II often found in the bibliography [228, 229, 233, 237, 273, 278, 281]. As discussed in chapter 5, there are several references looking to design aspects of this circuit, namely, dynamic range features [237, 278], noise performance [229, 281] and accuracy [273, 278, 281]. Frequency response optimization guidelines were also covered. The first contributions about translinear CC-II frequency optimization, arising from simulation and measured evidence were published in references [288, 289, 292] by the Nero Alves et al. Chapter 5

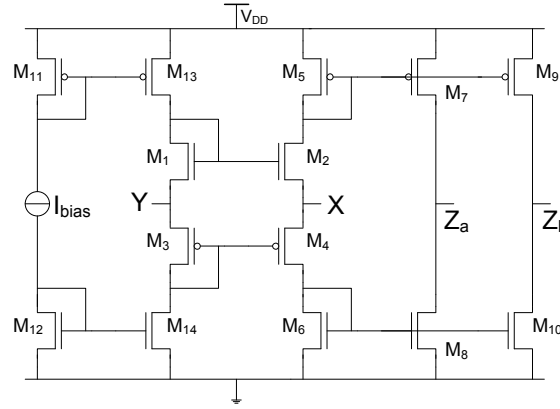


Figure 6.20 – Class AB translinear CC-II.

provided for the first time theoretic guidelines towards frequency response optimization for this special class of CC-IIs. The results reported on chapter 5 arise from a modeling strategy based on two aspects: frequency response normalization and level 2 MOSFET transistor modeling. Even though the detail of level 2 MOSFET models can not be compared with the high precision BSIM3v3 models often used in current CMOS technology design processes, the general theoretical guidelines of chapter 5 hold true and provide a good starting point for design optimization.

There are four fundamental guidelines disclosed from the theoretical framework of chapter 5 that need experimental demonstration, namely:

- Translinear CC-II using equalized transition frequency PMOS and NMOS transistors display very good cut-off frequency.
- The input impedance seen from the X input port is inversely dependent on the transconductance of the transistors.
- The current gain between input X and output Z is generally insensitive to device characteristics.
- The voltage gain between the Y input and the X output displays larger sensitivities to device characteristics and larger bandwidths than the current gain.

Several simulation tests were carried in order to demonstrate these guidelines, using the simulation set-up of Figure 6.21. This set-up includes three independent signal sources necessary to fully characterize the CC-II transfer functions. Additionally (not represented) both outputs Z_a and Z_b and input Y were biased with a fixed voltage of $1.65V$ ($V_{DD}/2$). The transistors aspect ratios were fully parameterized using NMOS transistor's design aspect as reference, thus:

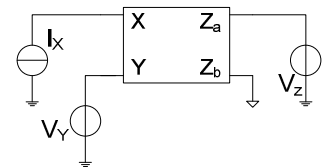


Figure 6.21 – CC-II optimization set-up.

- NMOS transistors M_1 , M_2 , M_6 , M_8 and M_{10} were designed with aspect ratio W_N/L and $L=350nm$.

- NMOS transistors M_{12} and M_{14} were designed with aspect ratio W_N/L_S and $L_S=1\mu m$.
- PMOS transistors M_3, M_4, M_5, M_7 and M_9 were designed with aspect ratio $\gamma W_N/L$ and $L=350nm$.
- PMOS transistors M_{11} and M_{13} were designed with aspect ratio $\gamma W_N/L_S$ and $L_S=1\mu m$.

Biasing transistors M_{11} to M_{14} used larger gate length transistors in order to have larger output conductance. All the remaining transistors used $L=350nm$ which is the minimum feature of the technology, thus assuring the least possible intrinsic capacitances for any possible W_N . The relevant design variables using this approach are limited to a number of three, W_N , γ and I_{bias} , making feasible a complete demonstration set.

The first task was to set the adequate value of parameter γ , the ratio between PMOS and NMOS aspect ratios. With $I_{bias}=100\mu A$ and $W_N=30\mu m$ several parametric simulations were carried out using γ as control parameter. Figure 6.22 shows the results of this analysis, using the current gain as reference. As can be seen, there is an optimum value of γ for which the current gain cut-off frequency attains its maximum value, lying between 0.3 and 0.4. This value was compared to the mobility ratio obtained from the transistor's model files. Since these transistors are modeled using BSIM3v3 specifications, the effective mobility (without velocity saturation and lateral field effects) was used as reference, holding $\mu_{p0}=148.2m^2/Vs$ and $\mu_{n0}=475.8m^2/Vs$. The mobility ratio becomes $\mu_{p0}/\mu_{n0}=0.311$, which agrees with these simulation results. This shows that, using equalized transition frequency transistors, with $\gamma=\mu_p/\mu_n$, leads to optimized cut-off frequency translinear CC-IIs. This experiment was repeated for other values of both I_{bias} and W_N showing that the optimum aspect was not dramatically affected. Figure 6.22 also shows that γ has negligible influence on the magnitude of the current gain, as predicted theoretically.

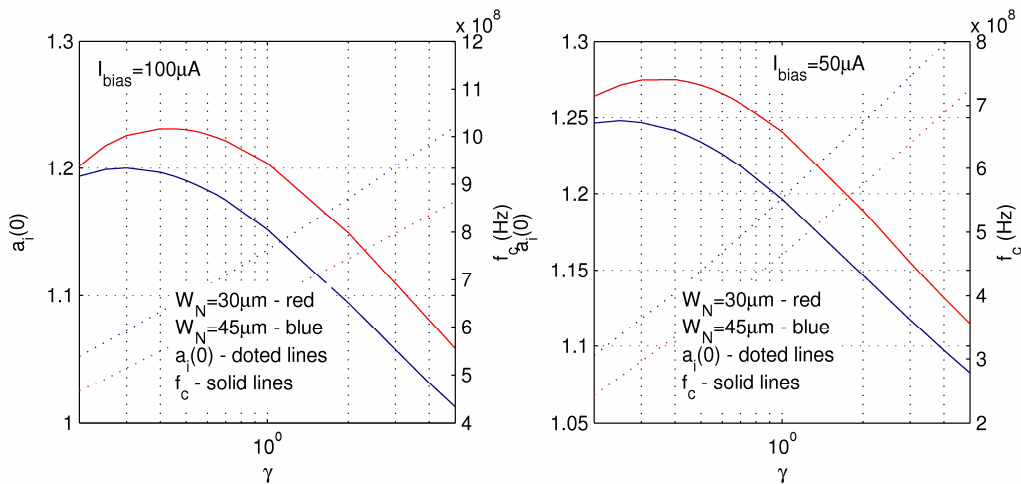


Figure 6.22 - γ optimization.

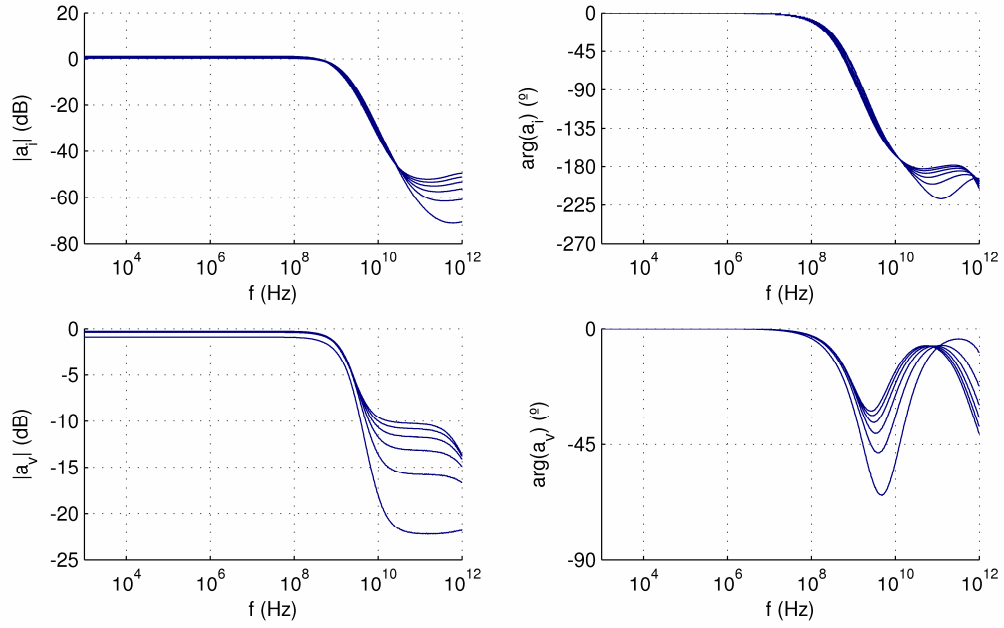


Figure 6.23 – Current and voltage gain frequency response for several values of W_N .

The following test investigated the influence of parameter W_N on the overall performance of the CC-II. Setting $I_{bias}=60\mu A$ and $\gamma=0.4$, several parametric simulation were carried out using W_N as control parameter. Figure 6.23, Figure 6.24 and Figure 6.25 show a qualitative representation of the effect of W_N on all the CC-II's transfer functions, for W_N ranging between $10\mu m$ and $60\mu m$. These figures should be compared with the correspondent theoretical prediction depicted on Figure 5.25, Figure 5.27 and figure 5.28 respectively. The fundamental aspects between theory and simulated results are in reasonable good agreement. Major deviations arise due to insufficient detail of the theoretical model adopted on chapter 5, which did not consider (for ease of treatment) several aspects, namely: i) the biasing current sources were ideal; ii) transistor models were limited to level 2 equations, suitable for hand analysis; and iii) body effect was not considered for the transistors forming the translinear loop. Nevertheless, the resemblance between theoretic and simulated results is remarkable, namely:

- The current gain curves agree well on magnitudes and pole contributions, displaying on both cases a maximum of 180° of phase shift.
- The Z driving point admittance is in complete agreement with the theoretical predictions.
- The Y driving point admittance differs from the theoretical perspective. Simulated results show that $y_y(s)$ displays a first order model arising from the parallel association of an equivalent resistance and a capacitance. Theoretical perspectives predicted a generally capacitive behavior. Simulation results showed that this behavior is attained if the bias current sources approach the ideal situation considered in the theoretical analysis.

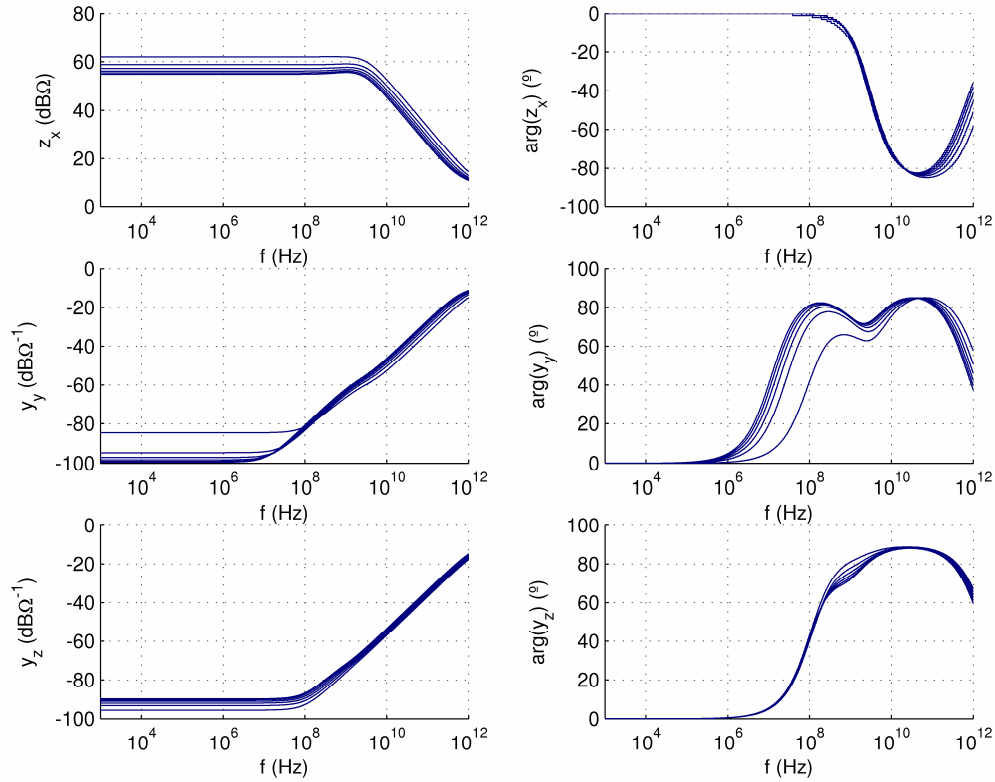


Figure 6.24 – Terminal driving point impedances (admittances) for several values of W_N .

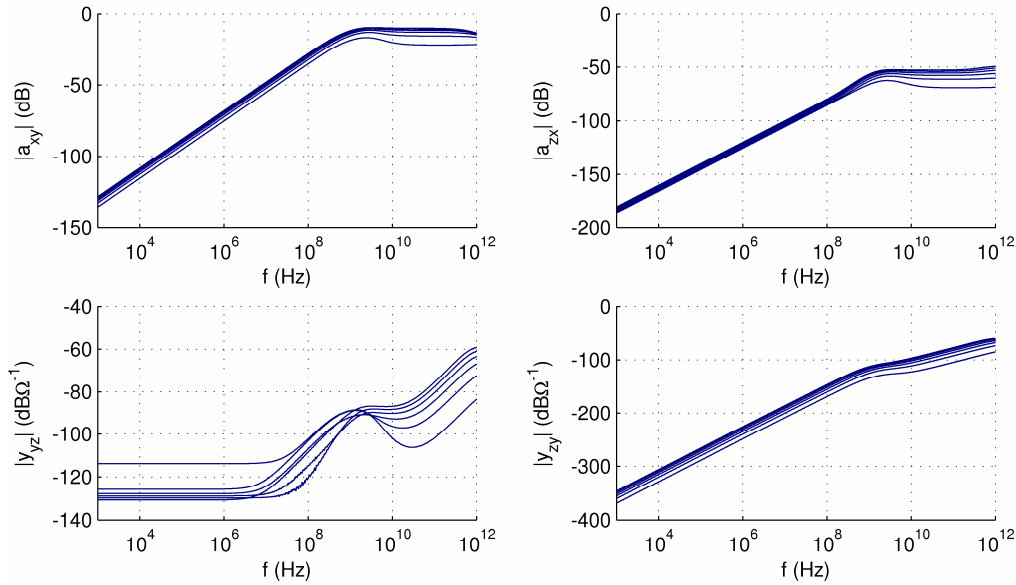


Figure 6.25 – Non-ideal transfer functions (reverse gains and reverse transfer admittances).

- Voltage gain curves show large variations both in magnitude and phase. This can be explained by body effect considerations that were not considered on theoretical modeling. Nevertheless, the voltage gain still shows larger bandwidths and less magnitude reduction.
- The X input impedance is in complete agreement with the theoretical perspectives, which mean, magnitude inversely dependent on W_N , and an approximate dominant pole

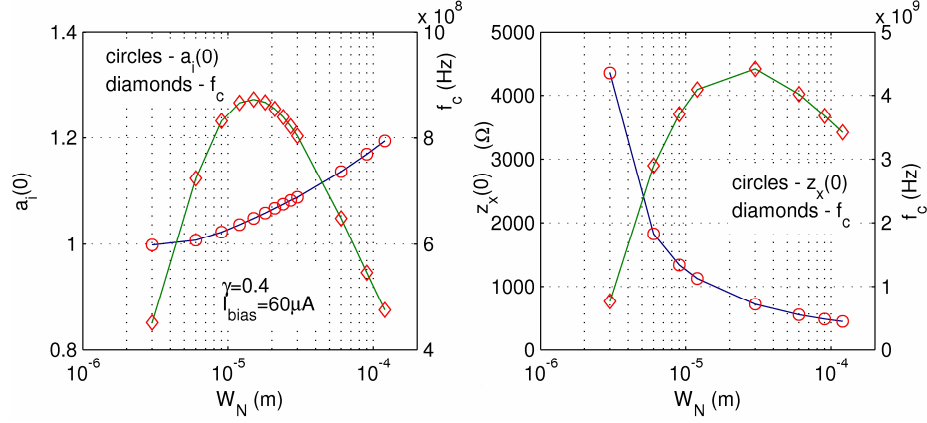


Figure 6.26 – Effect of W_N on current gain and X input impedance.

behavior with a maximum of 90° phase shift.

- Finally, the remaining non-ideal transfer functions (Figure 6.25) also displayed a reasonable good agreement with theoretical results. Since these transfer functions exhibit very small magnitudes both in theory and simulation perspectives, small deviations mean negligible impact on the overall performance.

Since both current gain and X input impedance are of special importance on CMD design, Figure 6.26 depicts the dependencies of both magnitude and cut-off frequencies of these transfer functions on W_N . Once again, the magnitude of the current gain exhibited small dependence on W_N . The cut-off frequency on the other hand exhibited a local maximum, thus meaning that for fixed γ and I_{bias} , W_N has an optimum value. The magnitude of the X input impedance decreases with W_N as predicted. The cut-off frequency is always larger than the cut-off frequency associated to $a_i(s)$. The plots of Figure 6.26 show that setting W_N to its optimum value is not a straight decision, since $z_x(0)$ can have great impact on the CMD performance. In this sense, a value of W_N corresponding to smaller impedance levels is obviously preferred.

Figure 6.27 shows the effect of I_{bias} on both $a_i(s)$ and $z_x(s)$. I_{bias} is certainly the parameter that has larger impact on CC-II characteristics. The magnitude of $a_i(s)$ exhibited nearly 35% variation for biasing currents ranging from $1\mu A$ to $30mA$. The cut-off frequency increases with I_{bias} for all observed values, approaching a limit value. As for $z_x(s)$, its magnitude decreases rapidly with I_{bias} and its cut-off frequency exhibited a maximum behavior. This maximum behavior is consistent with the theoretical predictions. The main reason for this arises from velocity saturation effects which occur with high current levels. Normally, it is not advisable to use such values of I_{bias} , since it would also imply large power consumption and restricted dynamic range (the V_{GS} of the MOS transistor increases with the drain current). Thus selecting an adequate I_{bias} is not often an easy task since it can compromise dramatically several aspects of the CC-II's performance.

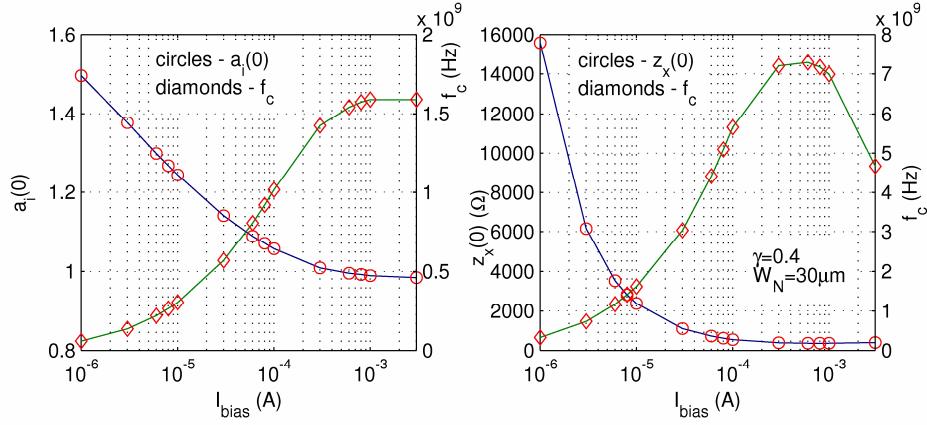


Figure 6.27 – Effect of I_{bias} on current gain and X input impedance.

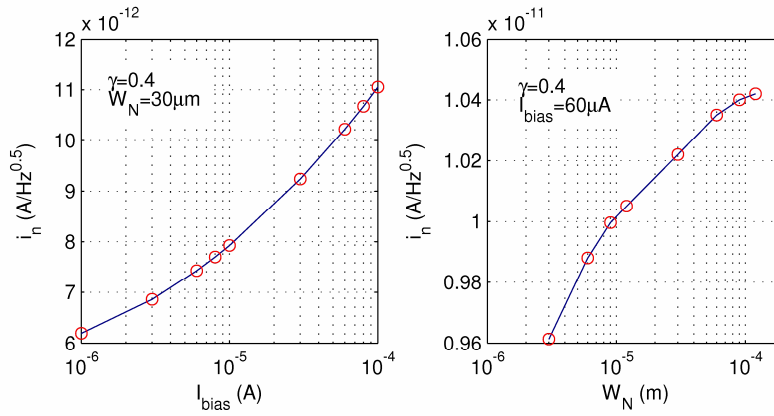


Figure 6.28 – Effect of I_{bias} and W_N on noise floor.

Finally, to complete the frequency response optimization study of translinear CC-II_s and also add some insight on the noise behavior of these circuits, some simulations were carried out. These simulations observed the impact of I_{bias} and W_N on the equivalent current input noise referenced to the X input. Figure 6.28 shows the results of these simulations concerning only the noise floor⁴⁰. As it can be seen, W_N has negligible impact on noise. For moderate values of W_N the noise floor increased less than 10%. However, it exhibited a strong dependence on I_{bias} , which poses another important performance trade-off.

Following all these simulations the values of γ , W_N and I_{bias} were set to 0.4, $30\mu m$ and $60\mu A$, respectively. These nominal values were chosen in order to provide the following characteristics:

- i) $z_x(0)$ smaller than $1k\Omega$. As it will be presented in the next section, higher values of $z_x(0)$ restrict both bandwidth and effectiveness of the controlling mechanism. For these

⁴⁰ From the noise minimization point of view it is of major importance to minimize the noise floor, since this minimum noise level occurs inside the amplifier's bandwidth, thus representing a useful approximation of the noise behavior of the amplifier.

parameters, $z_x(0)=720\Omega$.

- ii) The open-loop bandwidth characteristics of the CC-II (even without any source or load conditions), was also relevant for the CMD operation. As it was discussed on chapter 5 in a system comprising one CMD followed by one TIA stage, the pole positions have great influence on the overall bandwidth). For this design, $a_i(s)$ has a cut-off frequency of $800MHz$ which is much higher than the TIA second pole (near $200MHz$).
- iii) Power and area reduction were also important design issues. In order to proof its usefulness, the CMD technique must introduce the least possible power and area penalties. In this case the power consumption was $1.2mW$, and area was only $0.2nm^2$. Comparing with the power and area characteristics of the reference TIA (sections 6.3.1 and 6.3.2), these represented a power penalty of 11% and an area penalty of 5% .
- iv) On a less important scale was the equivalent input noise, which was around $10pAH_z^{-0.5}$, for the chosen parameter values.

Figure 6.29 shows the layout of the final CC-II. As it can be seen, the CC-II structure is highly symmetrical on both PMOS (upper row transistors) and NMOS (lower row transistors). Several dummy transistors were used to separate the CC-II's branches and also to provide good matching characteristics of the designed transistors.

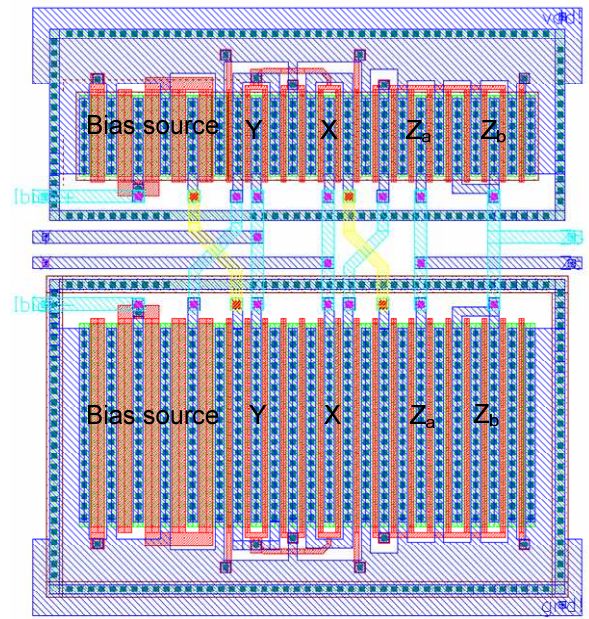


Figure 6.29 – CC-II layout.

6.4.3 CMD Design Optimization

This section presents the frequency response optimization procedure for an optical receiving system using only a CMD. Figure 6.30 shows the simulation set-up for this demonstration. As can be seen, the CMD consists on a multiple output CC-II with improved X input impedance. Following the results of chapter 5, the simplest feedback scheme for X input impedance reduction was adopted. This solution presents several advantages over the other techniques explored on chapter 5, namely: it has reduced complexity and it presents adequate

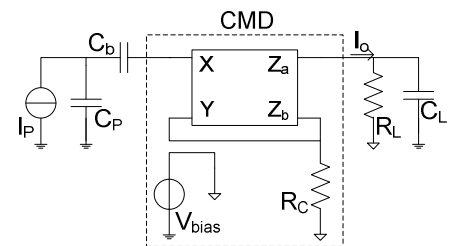


Figure 6.30 – CMD simulation set-up

solutions with reduced power and area penalty (since it only requires one CC-II plus some controlling elements). The simulation set-up also included the photo detector (a current source I_p in parallel with a capacitance C_P) and a load (representing the minimum conditions observed for the TIA experiment described on section 6.2, which are an input resistance of 200Ω (nearly the same as the TIA input resistance with $R_F=10k\Omega$) in parallel with a $6pF$ capacitance, necessary to assure stable operation of the TIA stage). A fixed voltage source, V_{bias} , establishes the biasing of all the CC-II ports. In order to avoid the use of bias blocking capacitors between CMD and TIA stage, this bias voltage was equal to the bias level of the TIA input transistor ($1.65V$).

As its was shown in chapter 5, this configuration suffers from the same bandwidth restrictions as a traditional TIA; the presence of a large capacitive source together with the amplifier's input impedance establishes a dominant pole. However, using the CMD circuit of Figure 6.30 and the controlling resistance R_C it is possible to lower this restriction. As it was discussed on chapter 5, this scheme employs positive feedback in order to reduce $z_x(s)$. The resultant $z_{xf}(s)$, (X input impedance with feedback) decreases linearly when R_C increases. The next simulation results provide evidence of this fact and also, how this affects the BWER of the configuration.

Figure 6.31 shows the frequency behavior of $z_{xf}(s)$ for several values of R_C . The CC-II used all the nominal parameters except W_N , which for this test took the value of $15\mu m^{41}$. The original open-loop impedance $z_x(s)$ had a maximum magnitude near to $60dB\Omega$, corresponding to $1k\Omega$. Actuating on R_C this maximum was reduced to nearly $30dB\Omega$ (near 30Ω). It is possible to reduce further the magnitude of the input impedance. It is even possible to simulate negative values of the input impedance. However, this should be used with special care, since the presence of a large capacitance at the input port imposes stringent stability restrictions.

Figure 6.31 also shows that $z_{xf}(s)$ exhibits large peaking phenomena. As it was explained on chapter 5, this results from the different cut-off frequencies between $z_x(s)$ and $a_i(s)$, which limit the feedback action. These peaks are limited by the open-loop impedance $z_x(s)$, displaying a maximum value which is slightly larger than $z_x(s)$ at the same frequency (which arise as a result from the non-idealities of the CC-II). The effect of these peaking phenomena is of small consequence for situations where C_P has large values and the input impedance reduction is moderate. For these situations the impedance imposed by C_P at the peak frequencies is small enough to attenuate the peaking effect. However, for situations where the input impedance reduction is severe, the peaking effect becomes noticeable even for high values of C_P . For these situations, the stability can be

⁴¹ W_N assumed several values for these simulations in order to fully observe the impact on both z_{xf} and BWER functions. These considerations justified the deviation of W_N from its nominal value of $30\mu m$.

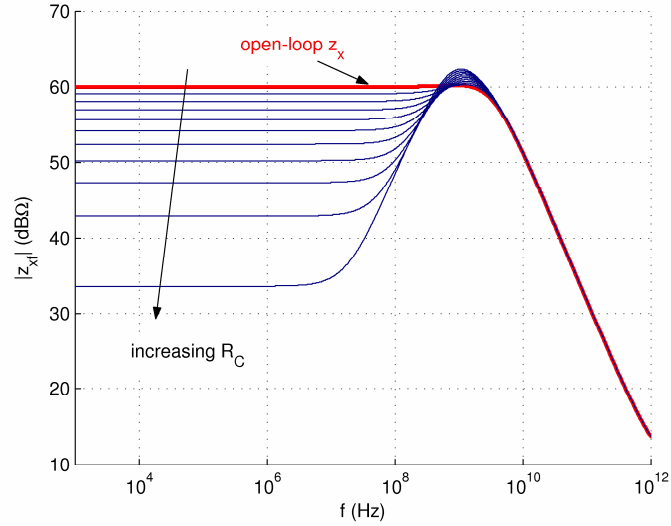


Figure 6.31 – X input impedance reduction.

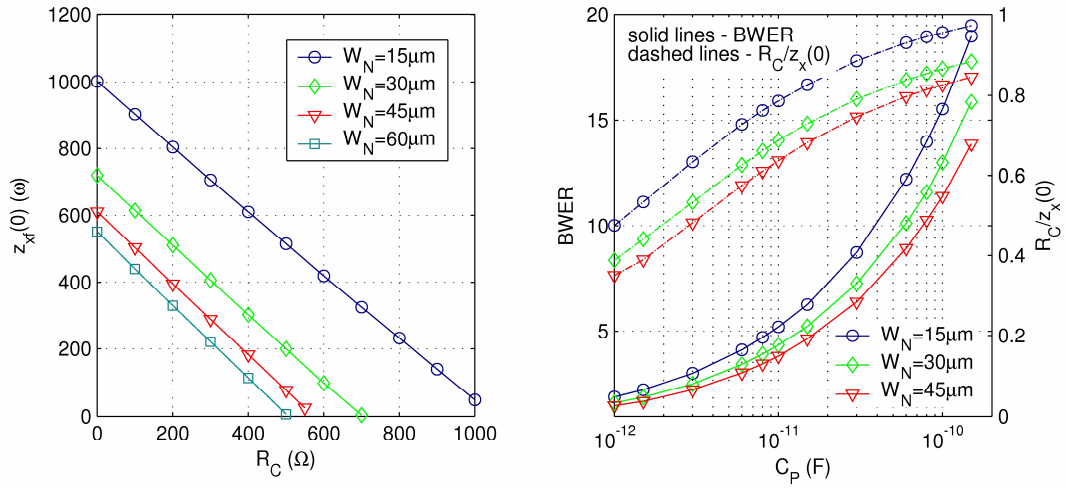


Figure 6.32 – Impedance control and BWER functions.

impaired.

Figure 6.32 shows the results of two simulation tests. The first test explored the impedance reduction effect. As it can be seen on Figure 6.32 (left plot), the impedance control follows a linear control depending on R_C . This study used several values of W_N in order to provide slightly different open-loop impedance conditions. Thus increasing W_N from $15\mu\text{m}$ to $60\mu\text{m}$ provide $z_x(0)$ values (the same as $z_{xf}(0)$ with $R_C=0$) ranging from $1k\Omega$ to 550Ω . As it is depicted on Figure 6.32, the starting impedance level has negligible effect on the linearity of the control scheme. Nevertheless, the control range on R_C reduces accordingly to $z_x(0)$; as it can be seen, $z_{xf}(0)$ attains almost the zero impedance condition for values of R_C approaching $z_x(0)$.

The BWER capabilities of this scheme were also investigated using the set-up of Figure 6.30. Several values of C_P ranging from $1pF$ to $150pF$ were used for this purpose. For each single value of C_P the current gain frequency response (taking I_p as input and I_o as output) was optimized using

R_C ; since the input pole is dominant for large values of C_P , using R_C to lower $z_{xy}(s)$ allows to improve bandwidth. The optimization procedure was based on the maximum flat frequency response criterion in order to provide an equal basis of comparison with the delayed feedback method. Figure 6.32 (right plot) shows the results of this optimization procedure. A maximum BWER improvement of 19 was achieved for a CMD consisting on a CC-II with $W_N=15\mu m$ and $C_P=150pF$. This represents a bandwidth enhancement from 1.06MHz of the original open-loop condition to 20.1MHz for the optimized version. These results show that the BWER attains large values for large C_P values, that is, for situations where the bandwidth is seriously restricted by the $C_P - z_x(s)$ interaction. Small values of C_P on the other hand exhibited small BWERs. This results from the fact that small values of C_P can not cope with the large frequency peaking present on $z_{xy}(s)$. Thus the impedance reduction is more conservative for this situations; this is also depicted on Figure 6.32 by the dashed lines showing the normalized $R_C/z_x(0)$ control values.

W_N has great impact on BWER. First of all, the BWER values decrease for increasing values of W_N . However this does not mean much about the effective bandwidth. It was observed that with larger values of W_N the bandwidth was higher for all values of C_P . The BWER reduction results from the fact that the reference open-loop bandwidth (with $R_C=0$) was also larger for these situations, implying a reduction of BWER. The final justification for $W_N=30\mu m$ arise from these considerations; with $C_P=150pF$ changing W_N from $15\mu m$ to $30\mu m$ increased the bandwidth from 20.1MHz to 23.4MHz. There was less gain for higher values of W_N (for instance, with $W_N=45\mu m$ the bandwidth was only 24MHz. Furthermore, with $C_P=1pF$ the bandwidth for these three situations was, 286MHz, 328.6MHz and 340MHz, respectively. So it seems justifiable to use $W_N=30\mu m$ instead of $15\mu m$, however, there is not much gain in changing it further to $45\mu m$.

6.4.4 CMD plus TIA Design Optimization

Figure 6.33 depicts the final simulation set-up, comprising a CMD and a TIA. C_P and I_P are used as before to model the photo-detector. C_b is only a bias blocking capacitor need to isolate the circuit's input from the photo-detector biasing circuitry. C_I represents the interstage capacitance. As it was discussed, the TIA stage is stable for input capacitance values larger than $6pF$. This was also the capacitance value taken as reference for the CMD design. Both CMD and TIA were design with the same biasing conditions. This design consideration avoids the usage of large blocking capacitances between CMD and TIA stages. Nevertheless, it should

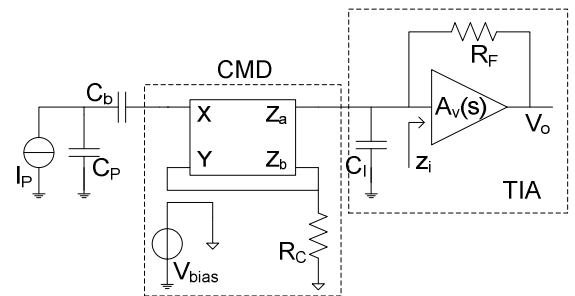


Figure 6.33 – TIA+CMD simulation set-up.

be remembered that since the interstage node is a high capacitive node ($6pF$ is a huge capacitance value for a $350nm$ process), it is possible to provide means of external access to this node. Thus, it would also be possible to use external bias blocking capacitors if desired.

A final version of the optimized CC-II, having $W_N=30\mu m$, $\gamma=0.4$ and $I_{bias}=60\mu A$ was designed using Cadence DFW - II. For this final circuit output Z_a providing the output current replica was designed with transistors NMOS and PMOS having $W_N=60\mu m$ and $W_P=24\mu m$, respectively. This decision was based on two aspects. First, the GBW of the final circuit depends on the gain of the configuration. In this sense, a gain of 2 is always better from the GBW point of view than 1. Second, from the noise performance point of view it is desirable to have an input stage with gain larger than unity, since this is the stage with largest noise contribution. Preliminary studies showed that the bandwidth loss on the CC-II stage due to gain improvements was only marginal for small gain factors.

The BWER performance of the TIA+CMD design was compared against two other similar conditions, both presented on section 6.3: i) the reference TIA (single TIA stage with no bandwidth optimization); and ii) the optimized TIA (single TIA stage optimized using delayed feedback). The simulation experiment for the TIA+CMD design perspective used the same values of C_P (ranging from $1pF$ to $150pF$) in order to provide a common basis for comparison means. TIA+CMD design optimization followed the same guidelines as before:

- Optimization based on post layout simulations in order to include the effect of parasitic capacitances.
- The TIA stage was first optimized using delayed feedback having $C_I=6pF$ present at its input. No further delayed feedback contribution was needed prior to this, since the interstage conditions do not change for the remaining situations.
- For each C_P value, the value of R_C was adjusted in order to provide maximum flat overall frequency response.

Figure 6.34 shows the bandwidth and BWER results achieved using the both delayed feedback and CMD concept. As it can be seen, the TIA+CMD design presents obvious advantages for both large and small values of C_P . For values of C_P less than $6pF$ it is not possible to use this TIA stage alone, since it would result in unstable behaviors. For very large values of C_P , the BWER provided by the TIA+CMD approach becomes competitive when compared to the other two reference possibilities (reference TIA and optimized TIA). However for values of C_P falling within the range from $6pF$ to $30pF$, delayed feedback (or even no means of optimization) presents superior performance. The reason for this behavior was discussed on chapter 5. For these situations the cut-off frequency of the TIA system alone is similar to the cut-off frequency of the CMD stage. This

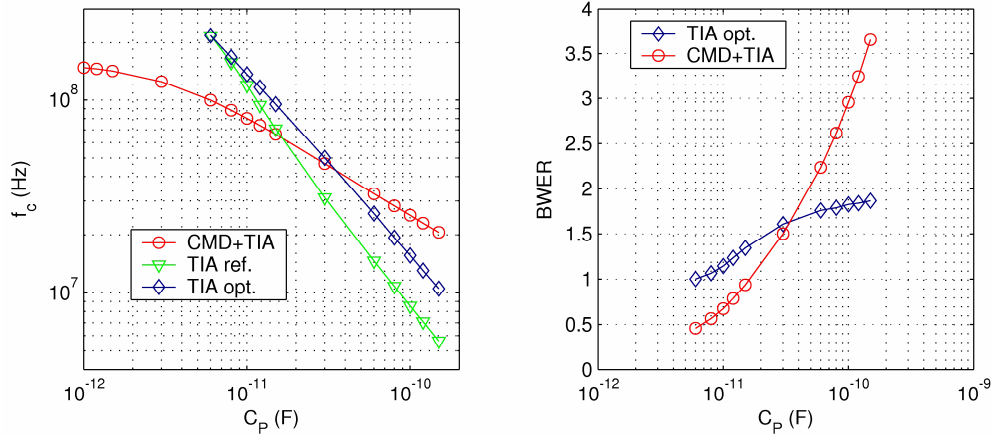


Figure 6.34 - Bandwidth and BWER behavior for several values of C_P .

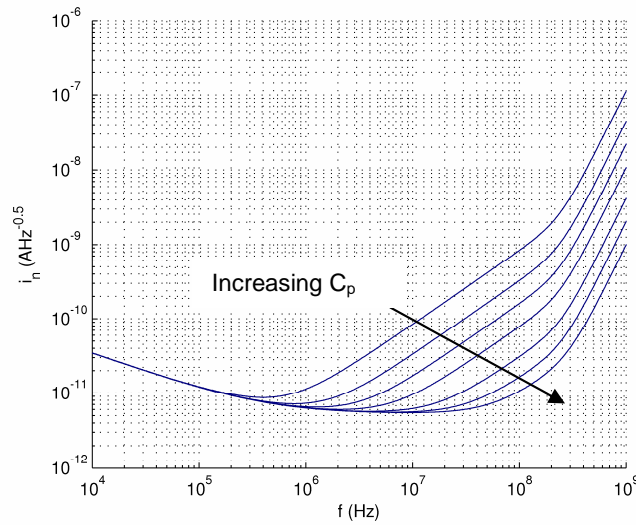


Figure 6.35 – Equivalent noise input current frequency behavior.

means that the dominant poles arising on the TIA and CMD stages are in the same frequency range, so the cascaded system presents bandwidth reduction. This fact also explains the bandwidth loss for the $C_P=1pF$; using a single CMD the bandwidth was around $320MHz$ for this situation. When the TIA stage (optimized for $C_I=6pF$ with a bandwidth $220MHz$) is added in cascade fashion, the overall bandwidth reduces to $150MHz$. The maximum BWER reached a value larger than 3.5, taking the reference TIA's bandwidth as reference, which meant a bandwidth improvement from $5.5MHz$ to $20.6MHz$. It is possible to achieve better results using these techniques if the bandwidths of the two stages are made very different.

Finally, to conclude the optimization process, some noise simulations were carried out in order to reveal the impact of the CMD stage on the overall noise performance. For this purpose, several values of C_P were considered. For each value of C_P the CMD was optimized following the above mentioned procedures. The equivalent noise input current displayed a minimum noise floor of $5.5pAHz^{-0.5}$ for the lowest values of C_P , as depicted on Figure 6.35. Compared to delayed feedback,

the CMD+TIA design strategy introduces superior noise penalty, since both noise floor and bandwidth increased.

6.5 Concluding Remarks

This chapter presented experimental evidence of the methods developed within this thesis framework. Three design examples were developed in order to demonstrate the validity of each of these methods. The first example addressed the experimental verification of the effects associated with the concept of delayed feedback. The following two examples addressed integrated circuit design issues using delayed feedback and active current matching. These last design examples followed a post layout simulation approach.

For the first design example, an experimental set-up consisting of standard discrete electronic components (TL082) was used to demonstrate the effects associated with the delayed feedback concept. The results presented in section 6.2 showed that:

- It is possible to achieve maximum flatness in feedback amplifiers comprising delay elements inside their feedback loops.
- There is a bandwidth improvement under the maximum flatness condition.
- Both maximum flatness and stability bound depend on the amount of delay and the internal dynamics of the amplifier (poles and gain).

This thesis presents delayed feedback as new a bandwidth enhancement method suited for feedback amplifiers. Section 6.3 presents an exploration of this point of view: the design of a transimpedance amplifier suited for FSO systems, using integrated circuit design tools. This design used the standard 3.3V, 350nm SiGe HBT/CMOS technology from AMS. As it was discussed in chapter 5, the major bandwidth limitation in these amplifiers arises at due to the input circuitry. The interaction of a high capacitance photo-detector with the amplifier's input impedance imposes a dominant pole behavior. For this example the design goals were a transimpedance gain larger than $80\text{dB}\Omega$ for a photo-detector having an intrinsic capacitance, C_P ranging from 10pF to 150pF . Under these conditions, the usage of the delayed feedback concept allowed significant bandwidth improvements (almost 100%) for large C_P values. Simulations results showed also that the noise performance was not dramatically impaired by this concept (the noise floor remained essentially fixed $2.2\text{pAHz}^{-0.5}$, with or without delay element). The total power consumption was 10.5mW .

Finally, a third simulation example applied active current matching to remove the input circuitry bandwidth limitation, of the previous case. The design used the same SiGe technology as before. Some measures of the overall performance for a system using a photo-detector with 150pF intrinsic capacitance were: $86.4\text{dB}\Omega$ transimpedance gain and 20MHz bandwidth (corresponding to a

bandwidth enhancement of 260%). The equivalent noise floor was under $5.5pAHz^{-0.5}$ and power consumption of $12mW$. Both noise and power penalties result for this case from the added circuitry needed to implement the CMD stage.

It was also verified that active current matching allows a large range of photo-detectors to be considered. This is because the transimpedance amplifiers had stability problems for small values of the photo-detector's intrinsic capacitance, which could be conveniently removed using active current matching. This was also the reason that impaired the combined usage of the two optimizing methods in this design example. Due to stability restrictions the reference transimpedance amplifier was compensated using the photo-detector's intrinsic capacitance. An equivalent capacitance was also present when using the active current matching approach, thus meaning that the transimpedance stage could not profit from delayed feedback, for this situation. Nevertheless, the value of the compensation capacitance was determined taking into consideration the pole frequencies, the return ratio and the internal delay of the transimpedance amplifier.

7 CONCLUSIONS

“Um lente de repente não passa a ser doutor um agente a reagente um actor a reactor Um infante num instante não passa ser senhor uma pulga a elefante um s’tôr a professor Uma mente de repente não passa a ser brilhante um indigente a ser gente um SG a gigante Um jumento num momento não passa a inteligente uma finta a ser um tento uma bochecha a presidente” – Luciano Barbosa.

Summary

This chapter summarizes the most relevant conclusions of this work. A general appreciation of the proposed methods for bandwidth enhancement is addressed, concerning both delayed feedback and active current matching concepts. The adequacy of these methods for the design of FSO receiving systems is clarified taking as reference the results discussed in chapter 6. Finally some guidelines for future work relating are also foreseen.

7.1 Final Discussion

This work presented and explored two new design concepts on the design of large gain and bandwidth amplifiers. The objective was to develop strategies suitable to bandwidth enhancement in integrated circuit amplifiers. The example of design high gain and bandwidth transimpedance amplifiers for free space optical networks was taken as the motivating reference. Two different and complementary lines of thought were explored. These two lines were: input impedance reduction – since input impedance acted as a restricting characteristic in these amplifiers; and bandwidth optimization in feedback amplifiers. The concept of active current matching was advanced as suitable method for input impedance reduction in transimpedance amplifiers for FSO receiving systems. The concept of delayed feedback was proposed as a new method for bandwidth enhancement in feedback amplifiers.

7.2 Conclusions

The main conclusions regarding the above mentioned methods can be divided into three categories, relating to both the current matching device technique and the delayed feedback concept. A third category establishes the link between these two techniques and presents the overall system design conclusions.

7.2.1 Active Current Matching

The active current matching technique was proposed to reduce the input impedance in transimpedance amplifiers. This technique is especially suited for the design of FSO receiving systems, employing large area photo-detectors. According to this technique, a current matching device (CMD) is used between the photo-detector and the transimpedance amplifier's input. The main concept was to convey the input current converted from optical into electrical domains by one photo-detector, from a medium of ideally low input impedance to a medium of moderate impedance (the input of the transimpedance amplifier). These CMDs are best designed using current-mode techniques, allowing the exploration of high bandwidth circuits. For this purpose special second generation current conveyors were used, employing feedback in order to further reduce their input impedance. The main aspects of this technique are:

- Effective isolation between photodetector and transimpedance amplifier – allowing gain and bandwidth of the last transimpedance stage, to be taken separately.
- Effective method to control the input impedance and hence the bandwidth of the input circuit.
- Moderate complexity of the involved circuits.

- Suitable to integrated circuit design, since it does not involve the usage of complex elements such as inductances.

This technique also has some drawbacks, namely:

- Inherent increase in power consumption, due to the added active electronics.
- Performance penalty in noise behavior, due to the fact that low gain amplifiers at the input of a cascaded amplifier are not a good choice for noise optimization.
- Inherent increase in circuit area.

Using this technique it was possible to enhance the overall system's bandwidth by 30% for an equivalent input impedance reduction.

7.2.2 Delayed Feedback Concept

The delayed feedback technique was proposed as an effective method to exploit significant (almost 120%) bandwidth improvements in feedback amplifiers. This technique is based on the presence of delay elements inside the feedback path of these amplifiers. Since 1948, that these delays are known and suitable design techniques to compensate their associated instability effects have been demonstrated. However, until the present work, their usage as an effective means of bandwidth improvement has not been reported. The major benefits of this new concept are:

- Possibility of improving the closed-loop bandwidth in 141% (near 120% for maximum flat operation).
- Adequate means of controlling bandwidth and stability without being prone to device mismatch effects.
- Suitable to integrated circuit design.
- Low added complexity.
- Low power penalties.

The major drawbacks are:

- Difficult to predict the system bandwidth without special purpose numerical analysis.
- Added design variables, resulting in added design complexity.
- Prone to active device delays in very high frequency applications.
- Increased circuit area, specially using microstrip or stripline transmission lines.

The usage of this delayed feedback concept can be compared to other techniques, for example frequency peaking techniques, revealing its superiority offer.

7.2.3 Delayed Feedback versus Active Current Matching

The final result of this investigation culminated on the design of a transimpedance amplifier

suitable for free space optical reception. Both delayed feedback and active current matching concepts were used for the optimization of the amplifier. A comparison between these two methods shows that:

- i) Active current matching shows superior performance for situations employing high intrinsic capacitance photo-detectors (C_P).
- ii) On the other hand, for moderate to low values of C_P , delayed feedback has larger BWER than active current matching.
- iii) From an integrated circuit implementation point of view, delayed feedback introduces less design penalties, concerning both area and power consumption issues.

This comparison suggests that these methods are applied in an independent fashion. There are situations where this is indeed the best design decision. However, it is possible to employ both methods to the optimization of the same transimpedance amplifier. As it was discussed in chapter 5, CMD acts outside the feedback loop of the TIA stage. Thus it is possible to extend the bandwidth of the TIA plus CMD system using: i) delayed feedback to optimize the TIA stage; and ii) active current matching to reduce the dominant pole due to the photo-detector intrinsic capacitance.

7.3 Guidelines for Future Work

This work raised several questions and left unfinished several design perspectives that can profit from further exploration. Among the unanswered questions the following four deserve future attention:

- The fundamental limitation of GBW in feedback amplifiers; there is no known result summarizing the general GBW limitations of feedback amplifiers, so there is much to gain in answering this question, and many new useful design approaches to explore. Bode's GBW limitations are applicable to interstage design. There is no closed solution for the GBW limitation of a chain amplifier, as discussed in section 2.3. Feedback amplifiers often comprise more than one transistorized gain stage, thus falling into this category. According to the discussion in section 2.7.2, it is possible to improve the bandwidth of feedback amplifiers adding more gain stages to the feedback loop. However, this does not provide limits for GBW or on the number of stages.
- Following the same fundamental line of thought, the question of the best achievable GBW in chain amplifiers having flat response is another unanswered question [25, 30]. Chain amplifiers can be optimized using maximum GBW interstage design, as was done in section 2.3. Nevertheless, it is possible to achieve similar results without posing any restriction on the type of interstages.

- The problem of whether or not dynamic distortion effects may constrain the usage of delayed feedback in amplifiers. The theory on delayed feedback amplifiers followed a small signal analyses approach. It is well known that feedback systems comprising delays have complex stability dynamics. This is worsened if these feedback systems may comprise also, non-linear elements. These non-linear elements can turn the response of the feedback system dependent on the signal amplitude. Several results reveal that such feedback systems may exhibit chaotic behaviors [282]. Thus it is of paramount importance to have a clear understanding on the conditions to assure stability.
- The problem of noise minimization in delayed feedback amplifiers was not addressed. The reason for this departs from traditional amplifier optimization strategies; it is often possible to optimize both noise and bandwidth in one amplifier without mutual impairments. In fact, noise and bandwidth optimization procedures can be made orthogonal [280]. Delayed feedback amplifier design presents new optimization methodologies that may not followed such simplifying assumptions. Thus it is necessary to establish noise minimization strategies for these amplifiers.

Among the unfinished design perspectives, the following can be of relevant importance:

- Developing efficient strategies to the design of maximal flat delay response in delayed feedback amplifiers. Chapter 3 was mainly focused on bandwidth enhancement using maximum flat frequency response arguments. It should be possible to optimize delayed feedback amplifiers in order to have maximum flat delay response.
- The design of delay elements using artificial distributed networks. Chapter 4 presented two suitable strategies for the design of delay elements, one using transmission lines and the other based on all-pass approximations of the ideal delay. Other possibilities may resort to the usage of artificial distributed networks to synthesize the behavior of a transmission line.

The first four problems present serious difficulties, since they rely on the development of an adequate theoretical framework. On the other hand, the final problems are confined to well known design approaches. Nevertheless, they can provide interesting research topics for further dissemination.

Appendix

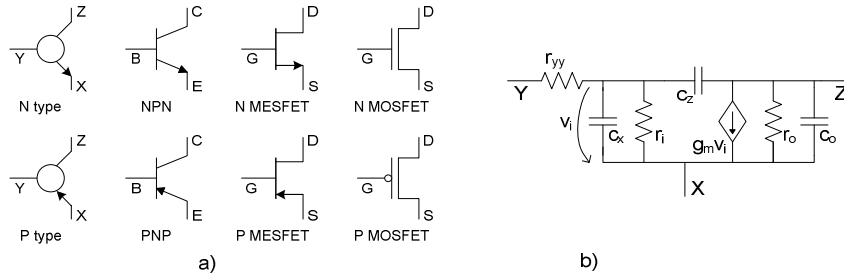


Figure A1.1 - Generic transistor: a) symbols, b) small-signal equivalent.

A 1 Generic Transistor

In this text a generic transistor model is used in order to make the discussion as general as possible. The usage of generic transistor models [267, 277, 280], abbreviates some considerations that should remain true, irrespectively to the type of active device in use (Bipolar, MOSFET, MESFET). When technology should become important, considerations toward the adaptation of the discussed methods can be made; this should be the case of special purpose very high frequency transistors, as HEMT and P-HEMT, for these cases the generic model may become to simplistic. Figure A1.1a, displays the adopted symbols for both N type and P type devices. A comparison with common transistor (Bipolar, MOSFET and MESFET) symbols is also present, to state the correspondence between them.

Although the difference between current-voltage characteristics of real transistors, the small signal equivalent circuit has several common elements. This observation allows the construction of a general model for the representation of both bipolar and field effect transistors small signal equivalent circuits; this model is presented on Figure A1.1b. The proposed model comprises a voltage controlled current source, g_m , output and input resistance, r_o and r_i respectively, intrinsic capacitances c_x and c_z , output parasitic capacitance c_o (normally, this last capacitance is neglected during circuit analysis, however, for a better comprehension of the transistor's frequency limitations this capacitance has to be accounted), and a Y terminal spread resistance r_{yy} . This last element together with c_o , has major impact in transistor performance at high frequencies of operation. Its impact can be greater than the loss resistances associated to the other contacts, because in general the Y terminal (Base for Bipolar devices and Gate for FET devices) is in general formed using lower conductance materials (typically polysilicon or similar compounds). In contrast, the other device contacts are made of high conductance materials (metals, as aluminum, copper or gold), with smaller associated loss resistances [267].

Generic Transistor	Bipolar Transistor	FET Transistor
g_m	g_m	g_m
r_o	r_o	r_o
r_i	r_π	∞
c_x	c_π	c_{gs}
c_z	c_μ	c_{gd}
c_o	c_c	c_{ds}
r_{yy}	r_{bb}	r_g

Table A1.1 - Element comparison.

The Table A1.1 establishes a comparison between Bipolar and FET devices, using the small signal model of Figure A1.1b as starting point.

For both Bipolar and HBT devices, this model is adequate to specify small signal behavior. A more detailed model should also include the interactions with the substrate (in general both Bipolar and HBT use VBIC simulation models that include these effects [165, 167, 170, 173]). For the case of FET transistors, this model is also sufficient; however some considerations should be made: i) MOSFET devices may display what is known as body effect [169]. This effect

traduces as an influence between different potentials between source and substrate on the effective threshold voltage. At a small signal level this is represented as another voltage controlled current source in parallel with g_m . This effect can be switched off if an appropriate connection between source and substrate is considered during design (not always possible); ii) MESFET devices usually use a more detailed input circuit description. In general this resort in the usage of a finite resistance, connected in series with a capacitance (giving DC infinite resistance) [162, 164, 166].

The transconductance parameter g_m , has a fixed value for moderate frequency operation. However for high frequency applications, excess phase terms can affect the accuracy of the models, demanding a better description of these effects. There are several approaches to this problem, and several fitting solutions are used to simulate the delay of the transistor [157-173] (the first transistor models include this excess phase term in the current gain parameter α [157-161]). In general this fitting is attached to the g_m element, and is dependent on the transit times associated with the current conduction mechanisms inside the device. For instance, for bipolar and HBT devices the excess phase term is in general modeled using such forms as,

$$\alpha(\omega) = \frac{\alpha_o}{1 + j\omega/\omega_b} \quad (\text{A } 1.1)$$

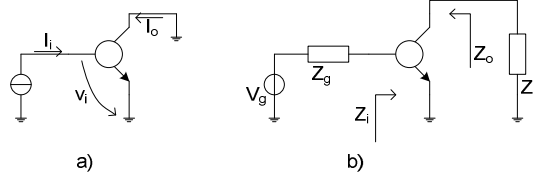


Figure A1.2 - Set-up for the definition of: a) f_T and, b) f_{max} .

$$\alpha(\omega) = \alpha_o \frac{\sin(\omega\tau_c)}{\omega\tau_c} e^{-j\omega\tau} \quad (A 1.2)$$

In (A 1.1) ω_b represents a 1st order approximation of the effect caused by the base-emitter delay. In (A 1.2) τ_c accounts for the collector transit time and τ is an empirical delay time consisting of the collector transit time plus a small portion of the base transit time.

Physical limitations of transistorized circuits are usually associated with transistor's figures of merit. Common figures are the transition frequency and the maximum oscillating frequency. The transition frequency is defined as the frequency at which the transistor current gain in a common X topology (see Figure A1.2a) reduces to 1. This is given approximately,

$$f_T \approx \frac{g_m}{2\pi(c_x + c_z)} \quad (A 1.3)$$

The ratio between g_m and the sum of the intrinsic capacitances is commonly referred as a limiting factor of the gain-bandwidth product in a wide variety of circuits. The transition frequency is not, however, independent of the circuit configuration. Another figure that has the property of being independent of the transistor configuration is f_{max} . This is the frequency at which the unilateral power gain (U) of the device becomes 1. U is the same as the maximum available power gain, taking the device as unilateral. It is possible to show that U is independent of the transistor configuration [94, 267]. U can be calculated using the configuration depicted on Figure A1.2b, assuming complex conjugated matching at both ports, that is $Z_i = Z_g^*$ and $Z_o = Z_L^*$. For a bipolar transistor [157, 158], under these considerations f_{max} is given approximately,

$$f_{max} \approx \sqrt{\frac{\alpha_o f_\alpha}{8\pi r_{yy} c_o}} \quad (A 1.4)$$

where, α_o represents the CY low frequency current gain ($\alpha_o = \beta_o / (1 + \beta_o) \approx 1$), and f_α represents the α cut-off frequency (also, the CY current gain [157]). Equation (A 1.4) shows that r_{yy} is indeed an important factor when the high frequency operation of the transistor is concerned. These two figures of merit are extensively referred all over the text. Evaluating the CY current gain, results,

$$\alpha(s) = \frac{\alpha_o}{1 + s(1 - \alpha_o)r_i c_x} \quad (A 1.5)$$

From (A 1.5) the α cut-off frequency is,

$$f_\alpha = \frac{1}{2\pi(1 - \alpha_o)r_i c_x} \quad (A 1.6)$$

Using (A 1.6) and the relation between g_m , β and r_i on a bipolar transistor, $\beta = g_m r_i$, f_{max} becomes [168],

$$f_{max} \approx \sqrt{\frac{f_T}{8\pi r_{yy} c_o} \left(1 + \frac{c_z}{c_x}\right)} \quad (A 1.7)$$

Equation (A 1.7) expresses the relation between f_{max} and f_T . Since in general $2\pi r_{yy} c_o$ is much smaller than $2\pi(c_x + c_z)/g_m$, the geometric mean between this factors (both present on (A 1.7)), is larger than f_T . The factor $(1 + c_z/c_x)/4$ is always less than 1, however the frequency relations lead in general to f_{max} values larger than f_T .

For high frequency FET devices (especially MESFET and HEMT), f_{max} is different form (A 1.4) and (A 1.7). The principal reason for this difference lies on the fact that FET devices have virtually infinite low frequency gate resistance. A suitable model to evaluate f_{max} must use a slightly different Y terminal representation than the one represented in Figure A1.1. This is normally accomplished using a series combination of a capacitance c_x (c_{gs}) and a small resistance r_i (r_{gs}) between the Y and X terminals, the other elements remain the unaltered. The resulting f_{max} is given by [162, 168, 172, 265],

$$f_{max} \approx \frac{g_m}{4\pi c_x} \sqrt{\frac{r_o}{r_i}} = \frac{f_T}{2} \left(1 + \frac{c_z}{c_x}\right) \sqrt{\frac{r_o}{r_i}} \quad (A 1.8)$$

Usually r_o/r_i is larger than 2, resulting in f_{max} being larger than f_T .

A 2 Butterworth Polynomials

The approximation problem plays an important role in the synthesis of network functions. It is often required to synthesize a network function (driving point impedance or transfer impedance functions), fulfilling some prescribed characteristics. Butterworth polynomials are a special kind of polynomials with the special characteristic of producing the maximally flat response over a prescribed bandwidth. This special characteristic makes Butterworth polynomials, the best means to achieve the best possible bandwidth. Butterworth polynomials are generated using the following generating function [251, 267],

$$A^2(\omega) = G(j\omega)G^*(j\omega) = \frac{1}{1+k^2(\omega)} \quad (\text{A 2.1})$$

where $k(\omega)$ is taken by ω^n , n being the polynomial order. Using $H(s)=1/G(s)$ and $s=j\omega$ (A 2.1) becomes,

$$|H(j\omega)|^2 - k^2(\omega) = 1 \quad (\text{A 2.2})$$

This is known as the Feldtkeller's equation, and it states that the loss ($H(s)$) approaches unity wherever the characteristic function $k(\omega)$ approaches 0. Using the loss function it is possible to explore the striking property of Butterworth polynomials. Taking the k^{th} derivative of (A 2.2) results in,

$$\frac{d^k |H(j\omega)|^2}{d\omega^k} = \frac{d^k k^2(\omega)}{d\omega^k} = \frac{(2n)!}{k!} \omega^{2n-k} \quad (\text{A 2.3})$$

The derivatives of the magnitude of the loss function have zero value at $\omega=0$ for $k \leq 2n-1$. Butterworth polynomials are the unique monic polynomials of order $2n$ having this property. This means that Butterworth polynomials are also the flattest possible near the origin. For this reason they are best known as maximally flat polynomials, leading to a maximally flat approximation procedure. Another interesting property is the constancy of cut-off frequency. In fact equating the cut-off frequency for the usual 3dB criterion, associated with (A 2.1) gives the normalized value of 1 rad/s.

The problem of finding realizable transfer functions able to produce the response in (A 2.1) is solved selecting out the roots on the LHP of $A^2(-s^2)$, since $G(s)G(-s)=A^2(-s^2)$. Equating $G(s)G(-s)$,

$$G(s)G(-s) = \frac{1}{1+(-1)^n s^{2n}} \quad (\text{A 2.4})$$

The poles in (A 2.4) are given by.

$$s^{2n} = (-1)^{n-1} \quad (\text{A 2.5})$$

Changing to polar coordinates, $s=\rho e^{j\phi}$,

$$\begin{cases} \rho = 1 \\ \phi = \frac{\pi}{2} + (2k-1)\frac{\pi}{2n} \end{cases} \quad (\text{A 2.6})$$

Where, k can take any integer value. Equation (A 2.6) proves one further interesting property: the roots of Butterworth polynomial lie on a circle of unitary radius. Selecting the roots with negative real part from (A 2.6) and factor them in a polynomial form gives for n even,

$$B_n(u) = \prod_{k=1}^{n/2} \left[1 + 2 \sin \left[\frac{(2k-1)\pi}{2n} \right] u + u^2 \right] \quad (\text{A 2.7})$$

and for n odd,

$$B_n(u) = (1+u) \prod_{k=1}^{(n-1)/2} \left[1 + 2 \sin \left[\frac{(2k-1)\pi}{2n} \right] u + u^2 \right] \quad (\text{A 2.8})$$

The variable u was used in (A 2.7) and (A 2.8) to symbolize the normalized low-pass approximation. Any other kind of characteristics can be approximated with Butterworth polynomials using an appropriate frequency change. For instance, s/ω_c and ω_c/s are used for low-pass and high-pass non-normalized approximations.

B 1 G_{max} Analytical Approximation

On section 3.1 it was stated that finding the cut-off frequency for delayed feedback amplifiers was a difficult task. The reason is that in general the solution of a quasi-polynomial is not expressible in an explicit form; however a wise manipulation of some of its terms may lead to important approximations. Starting with the cut-off frequency definition, $|H(j\omega)|^2 = 0.5|H(0)|^2$, simplifying some terms,

$$\frac{1+G^2}{(1+G)^2} + \frac{2G}{1+G} \left[\frac{\cos(\Phi)}{1+G} - K \sin(\Phi) \right] + K^2 = 2 \quad (B 1.1)$$

Assuming that the values of G are greater than 1 and that, as G increases further, some terms in (B 1.1) will become less and less significant and some others may approach unity, equation (B 1.1) can be approximated,

$$1 + K^2 - 2K \sin(\Phi) \approx 2 \quad (B 1.2)$$

Finally solving using the definition $\Phi = K(1+G)\delta$ in (B 1.2) results in,

$$1 + K^2 - 2K \sin(K(1+G)\delta) \approx 2 \quad (B 1.3)$$

Equation (B 1.3) is still difficult to solve explicitly, however it allows the investigation of the maximum value of K and the necessary condition on G . This result can be achieved by a simple investigation of the zero of the implicit derivative of K in order to G , followed by the proof that indeed this point, corresponds to a maximum of K (this conclusion is supported by the numerical results previously presented on Figure 3.4). Taking the derivative of (B 1.3) in order to G ,

$$\frac{\partial}{\partial G}(K) = \frac{K^2 \delta \cos(\Phi)}{K - \sin(\Phi) - \Phi \cos(\Phi)} \quad (B 1.4)$$

The first zero of (B 1.4) occurs when $\Phi = \pi/2$. It is clear that for this value the denominator of (B 1.4) is different from zero. Furthermore, the values of Φ that nullify (B 1.4) are infinite: it suffices to notice that every $\Phi = (1/2 + n)\pi$ (for integer values of n), is a solution of (B 1.4). Using $\Phi = \pi/2$ on (B 1.3) and solving the resulting quadratic equation in K , gives, $K = 1 \pm \sqrt{2}$. This result agrees with the predicted maximum according to Figure 3.4, also under the same condition; very large values of G . The negative result of K must be considered in absolute value. The ratio between two cut-off frequencies is always a positive value. Considering the lower limit on cut-off frequency, obtained using the lower bound function $H_{Lo}(\omega)$ (defined in equation (3.3)), results in $1/\sqrt{2}$. The second derivative of K can be obtained using equation (B 1.4) as starting point, and by evaluating its derivative for $\Phi = \pi/2$,

$$\left. \frac{\partial^2}{\partial G^2}(K) \right|_{\Phi=\frac{\pi}{2}} = -\frac{K^2 \delta}{K-1} \frac{\partial}{\partial G}(\Phi) = -(\delta)^2 \frac{K^3}{K-1} \quad (B 1.5)$$

Given the positive value of K previously calculated, the second derivative of K has a negative value for the first zero of (B 1.4), this justifies the assumption that indeed this is a local maximum of K . Finally, the necessary condition on loop-gain to achieve the first maximum of K can be easily calculated from the definition of Φ ,

$$G_{max} = \frac{\pi}{2(1+\sqrt{2})\delta} - 1 \quad (B 1.6)$$

A comparison between this value and the value G_{max} found numerically reveals that for $G > 1$ the approximation error is always less than 10%, which seems reasonable for design purposes.

B 2 Theorem 3.4 (Demonstration)

Theorem 3.4: A delayed feedback amplifier with loop transfer function $G(s)$ given by,

$$G(s) = G_o \frac{e^{-sL}}{1 + a_1 s}$$

where G_o , a_1 and L are all positive, is Hurwitz stable for G_o values satisfying the following restriction:

$$0 < G_o < \sqrt{1 + \frac{a_1^2}{L^2} z_1^2}$$

where z_1 is the solution of $\tan(z) = -\frac{a_1}{L} z$ in the interval $]\pi/2, \pi[$.

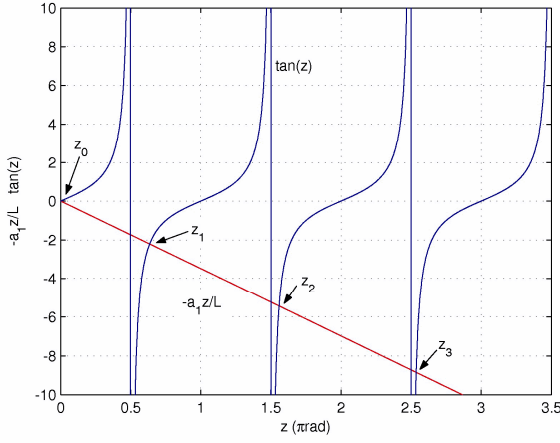


Figure B2.1 - Graphical solution of (B 2.6).

Let $\delta^*(s) = \delta(s)e^{sL}$ represent the characteristic quasi-polynomial, according to theorem 3.3. Using the frequency transformation $z = \omega L$ the real and imaginary parts of $\delta^*(j\omega)$ are expressed by,

$$\delta_r(z) = \cos(z) - \frac{a_1}{L} z \sin(z) + G_o \quad (\text{B 2.1})$$

$$\delta_i(z) = \sin(z) + \frac{a_1}{L} z \cos(z) \quad (\text{B 2.2})$$

Starting with the second condition of theorem 3.4,

$$\delta_i(\omega_0)\delta_r(\omega_0) - \delta_r(\omega_0)\delta_i(\omega_0) > 0 \quad (\text{B 2.3})$$

For some $\omega_0 \in]-\infty, \infty[$. Setting $\omega_0 = 0$, $z_0 = 0$. Thus $\delta_i(z_0) = 0$ and $\delta_r(z_0) = 1 + G_o$. Evaluating the derivative of $\delta_i(z)$ at z_0 , condition (B 2.2) becomes,

$$\delta_i'(z) = \left(1 + \frac{a_1}{L}\right) \cos(z) - \frac{a_1}{L} z \sin(z) \quad (\text{B 2.4})$$

$$(1 + a_1/L)(1 + G_o) > 0 \quad (\text{B 2.5})$$

Condition (B 2.5) is true under the assumed conditions, since both a_1 , L and G_o are positive.

The next step consists on the verification of the interlacing property of the roots of $\delta_i(z)$ and $\delta_r(z)$. Taking first the zeros of $\delta_i(z) = 0$, this results in the following non-linear equation,

$$\tan(z) = -\frac{a_1}{L} z \quad (\text{B 2.6})$$

The solution of (B 2.6) can not be expressed in a closed form. However, it is possible to depict a graphical solution of (B 2.6). According to Figure B2.1 the zeros of (B 2.6) can be found inside the interval sets,

$$z_0 = 0, \quad z_1 \in \left[\frac{\pi}{2}, \pi\right], \quad z_2 \in \left[\frac{3\pi}{2}, 2\pi\right] \dots$$

According to theorem 3.3, $\delta^*(s)$ has a principal term. Furthermore, $M=1$ and $N=1$. Taking $\eta = \pi/4$ and looking again to Figure B2.1, it is clear that inside $]0, 2\pi - \eta[$ $\delta_i(z)$ has exactly 3 zeros. Since $\delta_i(z)$ is an odd function it follows that inside $] -2\pi + \eta, 2\pi - \eta[$ there are exactly 5 zeros. Note also that $\delta_i(z)$ does not have any zero inside $]2\pi - \eta, 2\pi + \eta[$. Moreover, since $\delta_i(z)$ has two real roots in each of the intervals $[2l\pi + \eta, 2(l+1)\pi + \eta]$ and $[-2(l+1)\pi + \eta, -2l\pi + \eta]$, for $l=1, 2, \dots$. Hence $\delta_i(z)$ has exactly $4lN + M$ real roots inside $[-2l\pi + \eta, 2l\pi + \eta]$, which according to theorem 3.3 implies that $\delta_i(z)$ has only real roots.

Evaluating $\delta_r(z)$ at the roots of $\delta_i(z)$ it follows that, $\delta_r(z_0) = 1 + G_o > 0$. For the other zeros,

$$\delta_r(z_k) = \cos(z_k) - \frac{a_1}{L} z_k \sin(z_k) + G_o \quad (\text{B 2.7})$$

using (B 2.6)

$$\delta_r(z_k) = \frac{1}{\cos(z_k)} + G_o \quad (\text{B 2.8})$$

Using (B 2.6) it is possible to express $\cos(z_k)$ as,

$$\cos(z_k) = \frac{\pm 1}{\sqrt{1 + (a_1 z_k / L)^2}} \quad (\text{B 2.9})$$

Using (B 2.8) and (B 2.9) together with the distribution of z_k according to Figure B2.1,

$$\begin{aligned} \delta_r(z_k) &= G_o + (-1)^k \sqrt{1 + \left(\frac{a_1}{L} z_k\right)^2} \\ &= G_o + (-1)^k M(z_k) \end{aligned} \quad (\text{B 2.10})$$

The interlacing of roots is possible if the following set of conditions holds true,

$$(-1)^k \delta_r(z_k) > 0, \quad k = 0, 1, 2, \dots \quad (\text{B 2.11})$$

Since $M(z_k)$ is a monotonic increasing function of z_k for $z_k > 0$, and (B 2.11) is true for $k=0$, it suffices to verify the condition for $k=1$, that is,

$$G_o < M(z_1) \quad (\text{B 2.12})$$

This verifies the condition of theorem 3.4.

Finally, using the fact that the roots of $\delta_i(z)$ are all real and interlacing with the roots of $\delta_r(z)$, is sufficient to assure

that $\delta_r(z)$ has only real roots.

B 3 Theorem 3.5 (Demonstration)

Theorem 3.5: A delayed feedback amplifier with loop transfer function $G(s)$ given by

$$G(s) = G_o \frac{e^{-sL}}{1 + 2\xi s/\omega_n + (s/\omega_n)^2}$$

where G_o , L positive and Hurwitz stable denominator, is Hurwitz stable for G_o values satisfying the following restriction:

$$0 < G_o < \frac{\sqrt{(\omega_n L)^4 + 2G_o(\omega_n L)^2 z_1^2 + z_1^4}}{(\omega_n L)^2}$$

where z_1 is the solution of $\cot(z) = [z^2 - (\omega_n L)^2] / (2\xi\omega_n Lz)$ in the interval $]0, \pi[$.

The demonstration of theorem 3.5 follows the same steps of the previous demonstration. The assumptions for this case are, $G_o > 0$, $L > 0$, $\omega_n > 0$ and $\xi > 1/\sqrt{2}$. Starting with $\delta^*(s) = \delta(s)e^{sL}$ and performing the same frequency transformation as above results,

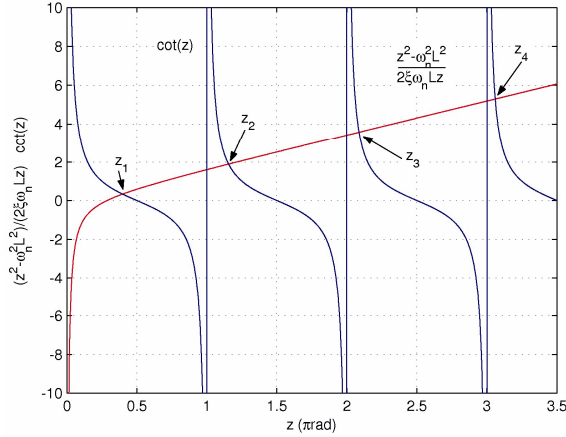


Figure B3.1 - Graphical solution of (B 3.2).

$$\delta_r(z) = \left(1 - \frac{z^2}{\omega_n^2 L^2}\right) \cos(z) - \frac{2\xi z}{\omega_n L} \sin(z) + G_o \quad (\text{B 3.1})$$

$$\delta_i(z) = \left(1 - \frac{z^2}{\omega_n^2 L^2}\right) \sin(z) + \frac{2\xi z}{\omega_n L} \cos(z)$$

Taking $z_0=0$, results $\delta_i(z_0)=0$ and $\delta_r(z_0)=1+G_o$. Evaluating the first derivative of $\delta_i(z)$ at z_0 , results in $\delta_i'(z_0)=1+2\xi/\omega_n L$, thus the second condition of theorem 3.3 holds true under the above assumptions.

Starting with the roots of $\delta_i(z)=0$, these roots can be expressed as the solutions of,

$$\cot(z) = \frac{z^2 - (\omega_n L)^2}{2\xi\omega_n Lz} \quad (\text{B 3.2})$$

These once again, can not be expressed in a closed form.

Figure B3.1 presents a graphical solution of (B 3.2). As can be seen, the solutions of (B 3.2) are distributed on the following intervals,

$$z_0 = 0, z_1 \in]0, \pi[, z_2 \in]\pi, 2\pi[..$$

According to theorem 3.3, $\delta^*(s)$ has a principal term. Furthermore, $M=2$ and $N=1$. Taking $\eta=\pi/4$ and looking again to Figure B3.1, it is clear that inside $]0, 2\pi-\eta[$ $\delta_i(z)$ has exactly 3 zeros. Since $\delta_i(z)$ is an odd function it follows that inside $]2\pi-\eta, 2\pi-\eta[$ there are exactly 5 zeros. Note also that $\delta_i(z)$ has zero inside $]2\pi-\eta, 2\pi-\eta[$, thus making a total of $4N+M=6$ zeros. Moreover, since $\delta_i(z)$ has two real roots in each of the intervals $[2l\pi+\eta, 2(l+1)\pi+\eta]$ and $[-2(l+1)\pi+\eta, -2l\pi+\eta]$, for $l=1,2,...$ Hence $\delta_i(z)$ has exactly $4lN+M$ real roots inside $[-2l\pi+\eta, 2l\pi+\eta]$, which according to theorem 3.3 implies that $\delta_i(z)$ has only real roots.

$z_0=0$ is also a solution of $\delta_i(z)=0$. The values of $\delta_r(z_k)$ are expressed by (B 3.1) taking $z=z_k$. Using (B 3.1) and (B 3.2) it is possible to write,

$$\delta_r(z_k) = G_o - \frac{2\xi}{\omega_n L} \frac{z_k}{\sin(z_k)} \quad (\text{B 3.3})$$

Solving (B 3.2) in order to $\sin(z_k)$ gives,

$$\sin(z_k) = \frac{\pm 2\xi\omega_n Lz_k}{\sqrt{(\omega_n L)^2 + 2(\omega_n L)^2 (2\xi^2 - 1)z_k^2 + z_k^4}} \quad (\text{B 3.4})$$

Knowing from Figure B3.1 the distribution of the zeros of (B 3.2), substituting (B 3.4) into (B 3.3) gives,

$$\delta_r(z_k) = G_o + (-1)^k M(z_k)$$

$$M(z_k) = \sqrt{1 + \frac{2}{(\omega_n L)^2} (2\xi^2 - 1) z_k^2 + \frac{z_k^4}{(\omega_n L)^4}} \quad (\text{B } 3.5)$$

In order to observe the interlacing property it is necessary to fulfill the following set of restrictions,

$$(-1)^k \delta_r(z_k) > 0, \quad k = 0, 1, 2.. \quad (\text{B } 3.6)$$

Since $M(z_k)$ is a strictly increasing function for $z_k > 0$ and $\xi < 1/\sqrt{2}$ conditions (B 3.6) are fulfilled if $G_o > 0$ and also (B 3.6) holds true for $k=1$, leading to,

$$G_o < M(z_1) \quad (\text{B } 3.6)$$

This verifies the condition of theorem 3.5.

Finally, using the fact that the roots of $\delta_i(z)$ are all real and interlacing with the roots of $\delta_r(z)$, is sufficient to assure that $\delta_i(z)$ has only real roots.

B 4 Mixed Delays – Stability Restrictions

Applying theorem 3.3 (the modified Hermite-Bihler Theorem for quasi-polynomials) for the case of mixed delays follows in a similar fashion as for the previous cases. However, this results in a set of non-linear equations with coupling parameters that can not be solved by procedures similar to the ones applied to find the roots of $\delta_i(z)=0$ in the previous cases. Starting with the loop-gain definition of equation (3.56), the characteristic equation is given by,

$$\delta(s) = (1 + a_1 s)(1 + Ts) + G_o(1 - Ts)e^{-sL} = 0 \quad (\text{B } 4.1)$$

It is clear that $\delta^*(s) = \delta(s)e^{sL}$ has exactly the same roots of (B 4.1). Separating the real and imaginary parts of $\delta^*(s)$, and using the frequency transformation $z = \omega L$,

$$\delta_r(z) = G_o + \left(1 - \frac{a_1 T}{L^2} z^2\right) \cos(z) - \frac{a_1 + T}{L} z \sin(z)$$

$$\delta_i(z) = \frac{G_o T}{L} z + \frac{a_1 + T}{L} z \cos(z) + \left(1 - \frac{a_1 T}{L^2} z^2\right) \sin(z) \quad (\text{B } 4.2)$$

The stability restrictions for (B 4.1) are expressed in two parameter space, where the parameters to consider are a_1/L and T/L . Comparing the $\delta_i(z)$ on (B 2.2) and (B 3.1) it is immediately apparent that both depend on the open-loop dynamical parameters except for G_o that does not appear in $\delta_i(z)$. In (B 4.2) both $\delta_i(z)$ and $\delta_r(z)$ exhibit the dependence on all the open-loop parameters, G_o included. This simple distinction inhibits the previous two step procedure: 1) find the roots of $\delta_i(z)=0$; 2) use this roots and $\delta_r(z)$ to verify the interlacing property – thus establishing the loop-gain stability constraints. The procedure to solve (B 4.2) is similar except, that now it is necessary to have a starting guess on G_o . This slight modification turns the procedure completely numerical. Similar procedures have been used to establish the stability bounds in PID controllers [150, 154, 155, 285].

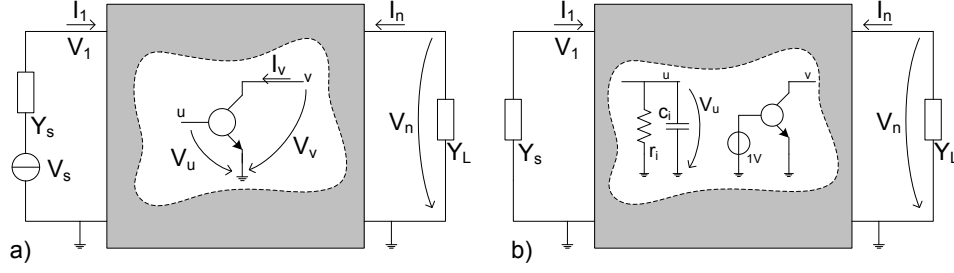


Figure C1.1 - Return ratio measurement: a) amplifier set-up; b) measurement procedure.

C 1 Measurement Procedures

The importance of accessing the return difference has been ascertained by Bode [248]. Since the stability properties of a given amplifier are intimately related with the zeros of its nodal admittance matrix (that must all lie inside the LHP), the definition of return difference of (4.6) shows that F has exactly the same zeros of Δ , provided that there is no cancellation of common factors between Δ and Δ^0 . Also, if Δ^0 is known to have no zeros on the RHP, which is usually the case for single loop feedback amplifiers, then F has exactly the same stability information as Δ .

There are several possible methods to measure the return ratio. The major hurdle in these methods is to measure F (or R) causing the least possible interference on the amplifier circuitry. As can be easily seen, using the β -A model comprises the tasks of measuring separately both the forward transfer and feedback transfer functions. Since the interaction between these blocks is complex for the majority of cases, this measuring procedure has to account for difficult loading effects. For this purpose, the feedback amplifier has to be adequately categorized into one of four possible configurations. These traditional configurations are: series-series, series-shunt, shunt-shunt and shunt-series; where the meaning of the words *series* and *shunt* refers to the combining and sampling methods used to return the feedback and measure the output signals, respectively. Despite of the particular feedback configuration in hand, these separation-categorization schemes usually result into invasive measuring procedures that may lead to erroneous results. Using the return difference instead provides a great simplification, since F can be measured in a fully independent fashion without having to identify the feedback configuration.

The most common method to measure the return difference is the Bleacher procedure [83]. By definition, F can be measured for any arbitrary element γ but preferably for a VCCS with one grounded terminal. One possibility is to choose any common X transistorized stage that may compose the feedback amplifier, and use its internal transconductance parameter as reference. Figure C1.1a shows the amplifier set-up, with one CX stage highlighted. Bleacher procedure consists in the following steps:

- 1) Killing all the independent signal sources that may feed the amplifier;
- 2) Breaking the Y terminal of the transistor;
- 3) Leaving attached to node u (the breaking point) the equivalent input impedance of the transistor (an adequate approximation for a CX stage is to use a parallel association of the input resistance r_i and the Miller equivalent capacitance $c_i = c_x + c_z(1 + g_m R)$); and
- 4) Drive the transistor with a unit voltage source as indicated on Figure C1.1b.

Under these conditions the return voltage is the negative of the voltage developing at node u (that is $R = -V_u/1V$). This procedure neglects the internal feedback of the transistor due to its intrinsic capacitance c_z . For the majority of the applications the Bleacher approximation of Figure C1.1b is sufficiently accurate.

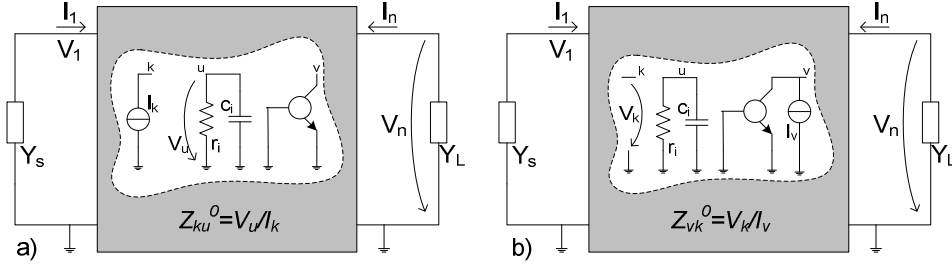
Other possible procedures specify the return difference as a ratio between two appropriate impedances [248]. Assuming the set-up of Figure C1.2a, for the transistor between nodes u and v , it is possible to choose any other circuit node, k , such that,

$$F = \frac{\Delta}{\Delta^0} = \frac{\Delta}{\Delta_{ku}} \frac{\Delta_{ku}}{\Delta^0} = \frac{Z_{ku}^0}{Z_{ku}} \quad (\text{C 1.1})$$

or

$$F = \frac{\Delta}{\Delta^0} = \frac{\Delta}{\Delta_{vk}} \frac{\Delta_{vk}}{\Delta^0} = \frac{Z_{vk}^0}{Z_{vk}} \quad (\text{C 1.2})$$

Since, Δ_{ku} and Δ_{vk} have no dependency on the transistor's transconductance g_m (the reference parameter, which appears into line and column u of the admittance matrix), they can both be interchanged with Δ_{ku}^0 and Δ_{vk}^0 respectively. This simplification step allows to express F as the ratio between two transfer impedances measured under normal operating conditions and with g_m set to zero (this means opening the Y terminal of the transistor). Z_{vk} and Z_{ku} represent the transfer impedances from node v to k and k to u , measured under normal operating conditions (respectively). Z_{vk}^0 and Z_{ku}^0 represent the same transfer impedances measured with g_m set to zero, according to the schemes of Figure C 1.2a) and b), respectively. Due to accessibility reasons node k can be chosen as the amplifier input or output node. Also,


 Figure C1.2 - Transfer impedance measurement: a) Z_{ku}^0 , b) Z_{vk}^0 .

the reference transistor can be adequately chosen as the first or the last transistor of the amplifier, for which $u=k$ or $v=k$, respectively. For these choices Z_{ku} and Z_{vk} become the driving point impedances Z_{uu} and Z_{vv} , measured at the input and output of the amplifier circuit.

Measuring the return ratio by Blecher's method or the return difference using impedance measurements are less invasive than the common loop-gain measurement of the β -A model. However, these are approximate methods since for both hold the assumption of negligible internal feedback on the reference transistor. It is possible to simulate the internal feedback using one extra transistor [83, 267], which represents a further complication that may not always be possible to handle. Fortunately it is possible to measure the return difference using two reference parameters [267]. Figure C1.3 represents the y parameter model of transistor with grounded X terminal. The driving point impedance seen from node u is a function of both y_{12} and y_{21} , since Δ is also a function of this two admittances, that is,

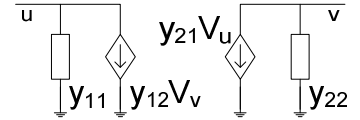


Figure C1.3 - Transistor y parameter model.

$$Z_{uu}(y_{12}, y_{21}) = \frac{\Delta_{uu}(y_{12}, y_{21})}{\Delta(y_{12}, y_{21})} \quad (C 1.3)$$

Assuming that, the primal interest is in measuring the return difference with reference to y_{21} , since for a transistor y_{21} is essentially given by the transistor's transconductance parameter, then according to the definition in (4.6),

$$F(y_{21}) = \frac{\Delta(y_{12}, y_{21})}{\Delta(y_{12}, 0)} \quad (C 1.4)$$

Multiplying equations (C 1.3) and (C 1.4),

$$F(y_{21})Z_{uu}(y_{12}, y_{21}) = \frac{\Delta_{uu}(y_{12}, y_{21})}{\Delta(y_{12}, 0)} \quad (C 1.5)$$

Since the elements y_{12} and y_{21} appear only once in row u and column u of the nodal matrix, Δ_{uu} has no dependency on both of these parameters, thus $\Delta_{uu}(y_{12}, y_{21})$ can be interchanged with $\Delta_{uu}(0, 0)$,

$$F(y_{21})Z_{uu}(y_{12}, y_{21}) = \frac{\Delta_{uu}(0, 0)}{\Delta(y_{12}, 0)} = \frac{\Delta_{uu}(0, 0)}{\Delta(0, 0)} \frac{\Delta(0, 0)}{\Delta(y_{12}, 0)} \quad (C 1.6)$$

The first quotient in the final branch of (C 1.6) represents the driving point impedance seen from terminal u when both y_{12} and y_{21} are set to zero, while the second represents the return difference with reference to y_{12} when y_{21} is taken zero. Thus $F(y_{21})$ can be measured using,

$$F(y_{21}) = \frac{Z_{uu}(0, 0)}{Z_{uu}(y_{12}, y_{21})} F(y_{12}) \Big|_{y_{21}=0} \quad (C 1.7)$$

The difference between (C 1.7) and the previous forms of (C 1.1) and (C 1.2) is that the impedance measurement without feedback is accomplished setting both y_{21} and y_{12} to zero, whether on the former cases only y_{21} was set to zero. To fulfill all the requisites of (C1.7) it is also necessary to measure one extra return difference. Figure C1.4 depicts the

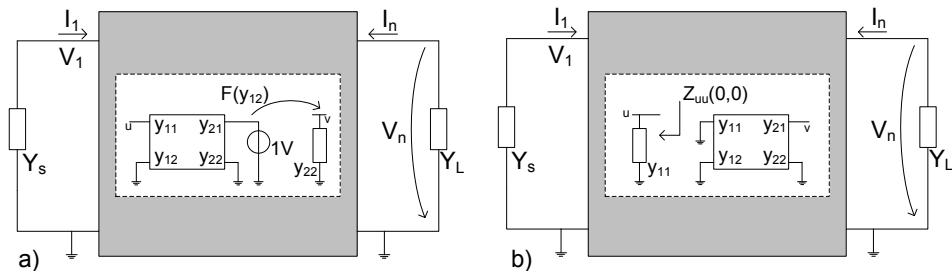


Figure C1.4 - Return difference measurement using driving point impedances.

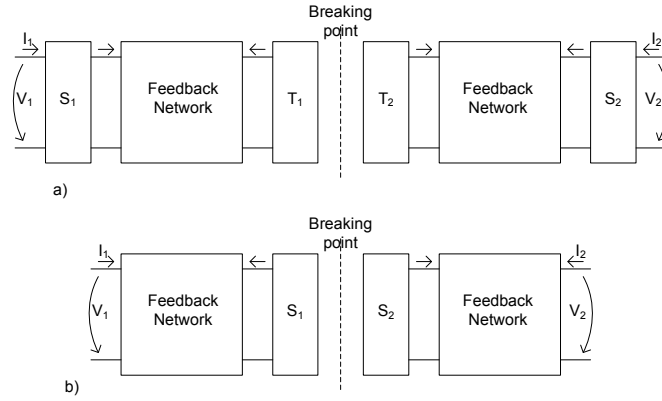


Figure C1.5 - Feedback network equivalent representations: a) external source model; b) internal source model.

conceptual procedures to measure the driving point impedance $Z_{in}(0, 0)$ (Figure C1.4b) and the return difference $F(y_{12})$ with y_{21} set to zero (Figure C1.4a). It is readily seen that the influence of y_{21} has been effectively removed on both schemes. Note that for the majority of cases, the return difference $F(y_{12})$ with y_{21} set zero is almost indistinguishable from unity and thus it is possible to evaluate only the two impedances in (C 1.7).

Both Bleecher's method and impedance measurements are suitable methods to measure the return ratio in real feedback circuits. Nevertheless, these methods have always some inherent simplifying assumptions that rend the final result inaccurate. In general, the inaccuracy comes from the fact that the active devices that compose the amplifier circuit may exhibit internal feedback. However, it is possible to devise exact measuring methods suitable for circuit design using CAD (Computer Aided Design) tools, as is the case of integrated circuit design.

Considering the case of a feedback amplifier comprising only linear and passive elements on the feedback network, one such possibility consists in substituting the feedback network by an adequate linear two-port model. Based on this approach Russel proposed the generalized internal source and external source models for feedback representation [105]. According to Russel's method, the feedback circuit can be converted into an external source or internal source equivalent model, as depicted on Figure C1.5. These equivalent representations replace the feedback circuit by two independent replicas using appropriate controlled source driving and termination conditions, inferred from the current to voltage relations of the original feedback network. The external source model adds two controlled sources (voltage or current) at end points of the original network, S_1 and S_2 in Figure C1.5a (the amplifier input and output). The breaking points are replaced by adequate short or open terminations, T_1 and T_2 . On the other hand, the internal source model adds two controlled sources at the breaking points (S_1 and S_2 on Figure C1.5b), thus simulating the closed-loop situation. Internal source model representations are usually more general, because the original bias conditions are preserved at both ends of the feedback network replicas. This makes, internal source model representations more adequate for feedback circuits comprising nonlinear elements. External source model representations are more adequate for linear feedback networks, since bias conditions are not affected by the short or open internal terminations.

Open-loop and closed-loop conditions are conveniently simulated using two circuit replicas: one for signal analysis and other for bias establishment. Selecting one of the controlled sources arising from the internal (or external) model representation as control parameter and its control signal as controlling variable, this resort to the following procedure:

- 1) Open-loop conditions are simulated setting the control signal equal to its replica on the bias circuit.
- 2) Closed-loop conditions are simulated restoring the control signal to its original position.

More details on the internal/external source models can be found in [105].

C 2 Blackman Theorem

Taking Figure C2.1 as the reference set-up, where the required driving point impedance at port X is under investigation. According to Thévenin's theorem, the equivalent impedance seen from any pair of terminals in a circuit is measured when all the independent sources are removed from the circuit (leaving their internal resistance in place) and substituted by an independent voltage source (can be a current source) V_T for test purposes. Then the required impedance is given as the ratio between the measured voltage and current at port X. When the circuit comprises feedback, more efficient methods resort to return ratio measurements and further calculations on simplified circuits (see for instance the measurement of direct and asymptotic gains for the Rosenstark

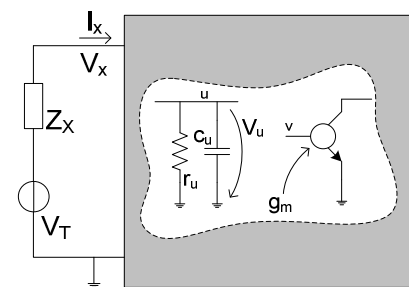


Figure C2.1 - Driving point impedance of a generic port X measurement set-up.

asymptotic model). Considering that the test source comprises internal impedance, Z_X , the return ratio for an arbitrary transistor placed at node v is measured taking Figure C2.1 as reference, with the transconductance parameter g_m as the control parameter and the voltage V_u as the control variable (under normal operation u and v are connected together). Clearly, since the circuit is assumed to be linear, this return ratio depends implicitly on Z_X ; this fact can be expressed by $R(Z_X)$, to reinforce the implicit relation. The Blackman theorem states that the driving point impedance seen from port X with feedback is given by the driving point impedance when the feedback loop is opened, multiplied by the ratio between the return difference measured with port X shorted against the return difference measured with port X opened, that is [78],

$$Z_{xf} = Z_{xol} \frac{1 + R(0)}{1 + R(\infty)} \quad (\text{C 2.1})$$

Equation (C 2.1) requires three independent calculations: the return ratio with port X shorted, $R(0)$; the return ratio with port X opened, $R(\infty)$; and the driving point impedance seen at port X when the control parameter is set to zero, Z_{xol} . The return ratio in both shorted and opened conditions can be calculated with port X terminated on a arbitrary test impedance Z_X , $R(Z_X)$, $R(0)$ and $R(\infty)$ are calculated taking the appropriate limits of $R(Z_X)$;

$$\begin{aligned} Z_{x,ol} &= Z_x \Big|_{g_m=0} \\ R(0) &= \lim_{Z_X \rightarrow 0} R(Z_X) \\ R(\infty) &= \lim_{Z_X \rightarrow \infty} R(Z_X) \end{aligned} \quad (\text{C 2.2})$$

Finally, it should be noted that the Blackman theorem represents a useful short-cut to the analysis of the effect of feedback on impedance. It is not apparent at first sight that three calculations are more efficient and less time consuming than going directly to the result. However, it should be reminded that from these calculations only one involves the complete circuit (the return ratio) and the other is done on a simplified circuit.

References

This section presents the list of the most relevant references relating to the described work. In order to allow an efficient and rapid way to consult the related references, this list was organized firstly into meaningful categories, according to the chapters content, and secondly by publishing date (inside each category).

FSO communication systems

1. F. R. Gfeller, U. Bapst, "Wireless In-House Data Communication via Diffuse Infrared Radiation", Proceedings of the IEEE, Vol. 67, November 1979.
2. C. Lomba, "Broadband Free-Space Atmospheric Optical Communications: A Design Study", Master dissertation, University College of North Wales, Bnagor, UK, November 1990.
3. J.. J. Puschell, R. Bayse, "High Data Rate Communication Systems for the Tactical Battlefield", IEEE Proceedings of the Tactical Communications Conference, Fort Wayne, Indiana USA, April 1990.
4. J. R. Barry, J. M. Kahn, E. A. Lee, D. G. Messerschmitt, "High-Speed Nondirective Optical Communication for Wireless Networks", IEEE Network Magazine, Vol. 5, November 1991.
5. R. L. Aguiar, "Implementação de um Sistema de Comunicação por Infravermelhos Utilizando Técnicas de Micro-Electrónica", Master dissertation University of Aveiro, Aveiro, Portugal, November 1994.
6. R. Valadas, "Redes de Comunicação de Área Local Não-Cabladas por Raios Infravermelhos", PhD Thesis, University of Aveiro, Aveiro, Portugal, November 1995.
7. A. Moreira, "Sistemas de Transmissão Ópticos em Espaço Livre para Ambientes Interiores", PhD Thesis, University of Aveiro, Aveiro, Portugal, February 1997.
8. J. M. Khan, J. R. Barry, "Wireless Infrared Communications", Proceedings of the IEEE, Vol. 85, February 1997.
9. C. Lomba, "Comunicações Ópticas em Espaço Livre para Ambientes Interiores: Modelação, Simulação e Optimização do Canal Óptico", PhD Thesis, University of Aveiro, Aveiro, Portugal, May 1997.
10. J. L. Cura, E. De Vasconcelos, R. L. Aguiar, "A High Transimpedance-Bandwidth, High Dynamic Range, CMOS Front-End for Open Air Optical Links", ICECS 98 – IEEE International Conference on Electronic Circuits and Systems, Lisbon, Portugal, September 1998.
11. E. de Vasconcelos, J. L. Cura, R. L. Aguiar, D. M. Santos, "A Novel High Gain, High Bnadwidth CMOS Differential Front-End for Wireless Optical Systems", ISCAS 99 – IEEE International Symposium on Circuits and Systems, Orlando Florida USA, June 1999.
12. R. L. Aguiar, A. Tavares, J. L. Cura, E. de Vasconcelos, L. N. Alves, R. Valadas, D. M. Santos, "Considerations on the design of transceivers for wireless optical LANs", IEE Electronics & Communications, Colloquium on Optical Wireless Communications, London, UK, June 1999.
13. G. Pang, K.-L. Ho, T. Kwan, E. Yang, "Visible Light Communication for Audio Systems", IEEE Transaction on Consumer Electronics, Vol. 45, November 1999.
14. L. N. Alves, R. L. Aguiar, E. De Vasconcelos, J. L. Cura, "A Sectorized Receiver for Infrared Wireless Networks", ISCAS 2000 – IEEE International Symposium on Circuits and Systems, Geneva, Switzerland, May 2000.
15. L. N. Alves, "Medição de Relação Sinal Ruído em Receptores Ópticos Atmosféricos", Master dissertation University of Aveiro, Aveiro, Portugal, November 2000.
16. L. N. Alves, J. L. Cura, R. L. Aguiar, D. M. Santos, "An Eight-Sector Wireless Infrared Receiver", ICECS 2000 – International Conference on Electronic Circuits and Systems, Jounieh, Lebanon, December 2000.
17. J. L. Cura, R. L. Aguiar, "Dynamic Range Boosting for Wireless Optical Receivers", ISCAS 2001 – International Symposium on Circuits and Systems, Sydney, Australia, May 2001.
18. M. Akanegawa, Y. Tanaka, M. Nakagawa, "Basic Study on Traffic Information Systems Using LED Traffic Lights", IEEE Transactions on Intelligent Transportation Systems, Vol. 2, December 2001.
19. G. Pang, T. Kwan, H. Liu, C.-H. Chan, "LED Wireless: A novel use of LEDs to transmit audio and digital signals", IEEE Industry Applications Magazine, Vol. 8, January 2002.

20. T. Komine, M. Nakagawa, "Integrated Systems of White LED Visible-Light Communications and Power-Line Communications", IEEE Transactions on Consumer Electronics, Vol. 49, February 2003.
21. T. Komine, M. Nakagawa, "Fundamental Analysis for Visible-Light Communications Systems using LED lights", IEEE Transactions on Consumer Electronics, Vol. 50, February 2004.
22. M. S. Shur, A. Zukauskas, "Solid-State Lighting: Toward Superior Illumination", Proceedings of the IEEE, Vol. 93, October 2005.

Theoretical GBW limitations

23. H. A. Wheeler, "Wide-Band Amplifiers for Television", Proceedings of the I.R.E, Vol. 27, July 1939.
24. W. W. Hansen, "On Maximum Gain-Bandwidth Product in Amplifiers", Journal of Applied Physics, Vol. 16, 1945.
25. R. M Fano, "Theoretical Limitations on the Broad Band Matching of Arbitrary Impedances", Research Laboratory of Electronics, Massachusetts Institute of Technology, Technical report No. 41, January 1948.
26. J. L. Stewart, Bandwidth Limitations in Equalizers and Transistor Output Circuits", I.R.E. Transactions on Circuit Theory, Vol. 4, March 1957.
27. J. J. Spilker, "Theoretical Limitations on the Gain-Bandwidth Product of Three-Terminal Networks", I.R.E. Transactions on Circuit Theory, Vol. 6, June 1959.
28. A. N. Perkins, "Theoretical Limitations on Wide-Band Multiterminal Networks", I.R.E. Transactions on Circuit Theory, Vol. 9, December 1962.
29. D. C. Youla, "A New Theory of Broad-Band Matching", IEEE Transactions on Circuit Theory, Vol. 11, March 1964.

Chain amplifiers

30. J. G. Linvill, "Amplifiers with Prescribed Frequency Characteristics and Arbitrary Bandwidth", Research Laboratory of Electronics, Massachusetts Institute of Technology, Technical report No. 163, July 1950.
31. B. Analui, S.-A. Hajimiri, "Method and Apparatus for a Multi-Pole Bandwidth Enhancement Technique for Wideband Amplification", US Patent No. 6778017, August 2004.
32. B. Analui, S.-A. Hajimiri, "Bandwidth Enhancement for Transimpedance Amplifiers", IEEE Journal of Solid-State Circuits, Vol. 39, August 2004.
33. Z. Lu, K. S. Yeo, J. Ma, M. A. Do, W. M. Lim, X. Chen, "Broad-Band Design Techniques for Transimpedance Amplifiers", IEEE Transactions on Circuits and Systems – I: Regular Papers, Vol. 54, March 2007.

Stagger tuned amplifiers

34. H. Wallman, "Stagger Tuned Amplifier", US Patent No. 2451893, October 1948.
35. H. Wallman, "Stagger Damped Tuned Amplifiers", US Patent No. 2480205, August 1949.
36. K. Honjo, Y. Takayama, "GaAs FET Ultrabroad-Band Amplifiers for Gbit/s Data Rate Systems", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-29, July 1981.
37. K. Honjo, T. Sugiura, H. Itoh, "Ultra-Broad-Band GaAs Monolithic Amplifier", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-30, July 1982.
38. K. T. Ng, B. Rejaei, J. N. Burghartz, "Substrate Effects in Monolithic RF Transformers on Silicon", IEEE Transactions on Microwave Theory and Techniques, Vol. 50, January 2002.
39. F. Carrara, A. Italia, E. Ragonese, G. Palmisano, "Design Methodology for the Optimization of Transformer-Loaded RF Circuits", IEEE Transactions on Microwave Theory and Techniques, Vol. 53, April 2006.
40. T. Biondi, A. Scuderi, E. Ragonese, G. Palmisano, "Analysis and Modeling of Layout Scaling in Silicon Integrated Stacked Transformers", IEEE Transactions on Microwave Theory and Techniques, Vol. 54, May 2006.
41. S. Shekhar, X. Li, D. J. Allstot, "A CMOS 3.1-10.6 GHz UWB LNA Employing Stagger-Compensated

Series Peaking”, IEEE Proceedings of the Radio Frequency Integrated Circuits Symposium, June 2006.

Distributed amplifiers

42. W. S. Percival, “Improvements in and relating to Thermionic Valve Circuits”, British Patent No. 460562, January 1937.
43. E. L. Ginzton, W. R. Hewlett, J. H. Jasberg, J. D. Noe, “Distributed Amplification”, Proceedings of the I. R. E., Vol. 36, August 1948.
44. W. H. Horton, J. H. Jasberg, J. D. Noe, “Distributed Amplifiers: Practical Considerations and Experimental Results”, Proceedings of the I. R. E., Vol. 38, July 1950.
45. D. V. Payne, “Distributed Amplifier Theory”, Proceedings of the I. R. E., Vol. 41, June 1953.
46. R. W. Scarr, “Discussion on: Distributed Amplifier Theory”, Proceedings of the I. R. E., Vol. 42, March 1954.
47. W.-K. Chen, “Theory and Design of Transistor Distributed Amplifiers”, IEEE Journal of Solid State Circuits, Vol. SC-3, June 1968.

Inductive peaking

48. A. Papoulis, “A new Class of Filters”, IRE National Convention, 1958.
49. M. Fukada, “Optimum Filters of Even Orders with Monotonic Response”, IRE Transactions on Circuit Theory, Vol. CT-6, September 1959.
50. J. F. O’Neill, M. S. Ghausi, “A Class of Realizable Maximally Flat Magnitude Driving-Point Functions”, IEEE Transactions on Circuit Theory, Vol. 14, March 1967.
51. N. Ohkawa, “Fiber-Optic Multigigabit GaAs MIC Front-End Circuit with Inductor Peaking”, IEEE Journal of Lightwave Technology, Vol. 6, November 1988.
52. J. Choma Jr., “Simplified Design Guidelines for Dominant Pole Amplifiers Peaked Actively by Emitter or Source Followers”, IEEE Transactions on Circuits and Systems, Vol. 36, July 1989.
53. J. J. Morikuni, S.-M. Kang, “An Analysis of Inductive Peaking in Photoreceiver Design”, IEEE Journal of Lightwave Technology, Vol. 10, October 1992.
54. I. M. Filanovsky, “A Generalization of Filters with Monotonic Amplitude-Frequency Response”, IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 46, November 1999.
55. S. S. Mohan, M. M. Hershenson, S. P. Boyd, T. H. Lee, “Bandwidth Extension in CMOS with Optimized On-Chip Inductors”, IEEE Journal of Solid-State Circuits, Vol. 35, March 2000.
56. W. Kossel, C. Menolfi, T. Morf, M. Schmatz, T. Toifl, “Wideband CMOS Transimpedance Amplifier”, IEE Electronics Letters, Vol. 39, April 2003.
57. S. Galal, B. Razavi, “Broadband ESD Protection Circuits in CMOS Technology”, IEEE Journal of Solid-State Circuits, Vol. 38, December 2003.
58. S. Galal, B. Razavi, “10-Gb/s Limiting Amplifier and Laser/Modulator Driver in 0.18- μ m CMOS Technology”, IEEE Journal of Solid-State Circuits, Vol. 38, December 2003.
59. C. Lee, C.-H. Wu, S.-I. Liu, “A 1.2V, 18mW, 10Gb/s SiGe Transimpedance Amplifier”, IEEE Asia-Pacific Conference on Advanced System Integrated Circuits, Fukuoka, Japan, August 2004.
60. S. Galal, B. Razavi, “40-Gb/s Amplifier and ESD Protection Circuit in 0.18- μ m CMOS Technology”, IEEE Journal of Solid-State Circuits, Vol. 39, December 2004.
61. C.-H. Wu, C.-H. Lee, W.-S. Chen, S.-I. Liu, “CMOS Wideband Amplifiers Using Multiple Inductive-Series Peaking Technique”, IEEE Journal of Solid-State Circuits, Vol. 40, February 2005.
62. S. M. R. Hasan, “Design of a Low-Power 3.5 GHz Broad-Band CMOS Transimpedance Amplifier for Optical Transceivers”, IEEE Transactions on Circuits and Systems – I, Vol. 52, June 2005.
63. S. Shekhar, J. S. Walling, D. J. Allstot, “Bandwidth Extension Techniques for CMOS Amplifiers”, IEEE Journal of Solid-State Circuits, Vol. 41, November 2006.

Capacitive peaking

64. J. A. Mataya, G. W. Haines, S. B. Marshall, "IF Amplifier Using Cc Compensated Transistors", IEEE Journal of Solid-State Circuits, Vol. SC-3, December 1968.
65. B. K. Ahuja, "An Improved Frequency Compensation Technique for CMOS Operational Amplifiers", IEEE Journal of Solid-State Circuits, Vol. SC-18, December 1983.
66. T. Wakimoto, Y. Akazawa, "A Low-Power Wide-Band Amplifier Using a New Parasitic Capacitance Compensation Technique", IEEE Journal of Solid-State Circuits, Vol. 25, February 1990.
67. G. Palmisano, G. Palumbo, "An Optimized Compensation Strategy for Two-Stage CMOS OP AMPS", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 42, March 1995.
68. G. Palmisano, G. Palumbo, "A Compensation Strategy for Two-Stage CMOS Opamps Based on Current Buffer", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 44, March 1997.
69. G. Palmisano, G. Palumbo, "Analysis and Compensation of Two-Pole Amplifiers with a Pole-Zero Doublet", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 46, July 1999.
70. F.-T. Chien, Y.-J. Chan, "Bandwidth Enhancement of Transimpedance Amplifier by a Capacitive-Peaking Design", IEEE Journal of Solid-State Circuits, Vol. 34, August 1999.
71. C.-W. Kuo, C.-C. Hsiao, S.-C. Yang, Y.-J. Chan, "2Gbit/S Transimpedance Amplifier Fabricated by 0.35 μ m CMOS Technologies", IEE Electronics Letters, Vol. 37, September 2001.
72. Y.-J. Chan, F.-T. Chien, T.-T. Shin, W.-J. Ho, "Bandwidth Enhancement of Transimpedance Amplifier by Capacitive Peaking Design", US Patent No. 6353366, March 2002.
73. J. Ramos, M. S. J. Steyaert, "Positive Feedback Frequency Compensation for Low-Voltage Low-Power Three-Stage Amplifier", IEEE Transactions on Circuits and Systems – I: Regular Papers, Vol. 51, October 2004.
74. J. Mahattanakull, "Design Procedure for Two-Stage CMOS Operational Amplifiers Employing Current Buffer", IEEE Transactions on Circuits and Systems – II: Express Briefs, Vol. 52, November 2005.

Feedback Amplifiers

75. H. S. Black, "Wave Translation System", US Patent No. 2102671, December 1937.
76. H. W. Bode, "Amplifier", US Patent No. 2123178, July 1938.
77. R. C. Shaw, "Wave Translation System", US Patent No. 2210503, August 1940.
78. R. B. Blackman, "Effect of Feedback on Impedance", The Bell System Technical Journal, Vol. XXII, October 1943.
79. S. J. Mason, "Feedback Theory – Some Properties of Signal Flow Graphs", Proceedings of IRE, Vol. 41, September 1953.
80. Y.-C. Ho, "Time-Domain Compensation for Closed-Loop Systems By a Delay Line Method", Research Laboratory of Electronics, Massachusetts Institute of Technology, Technical Report 288, January 1955.
81. S. J. Mason, "Feedback Theory – Further Properties of Signal Flow Graphs", Proceedings of IRE, Vol. 44, July 1956.
82. S. J. Mason, "Topological Analysis of Linear Nonreciprocal Networks", Proceedings of IRE, Vol. 45, June 1957.
83. F. H. Blecher, "Design Principles for Single Loop Transistor Feedback Amplifiers", IRE Transactions on Circuit Theory, Vol. CT-4, September 1957.
84. M. S. Ghausi, D. O. Pederson, "A New Design Approach for Feedback Amplifiers", IRE Transactions on Circuit Theory, Vol. CT-8, September 1961.
85. M. S. Ghausi, "Optimum Design of the Shunt-Series Feedback Pair with a Maximally Flat Magnitude Response", IRE Transactions on Circuit Theory, Vol. CT-8, December 1961.
86. E. K. Reedy, J. F. Pierce, "A Complementary Driver with a Subnanosecond Rise time", IEEE Journal of Solid-State Circuits, Vol. 4, October 1969.
87. B. L. Cochrun, A. Grabel, "A Method for the Determination of the Transfer function of Electronic Circuits", IEEE Trans. on Circuit Theory, Vol. CT-20, January 1973.
88. S. Rosenstark, "A Simplified Method of Feedback Amplifier Analysis", IEEE Transactions on

- Education, Vol. 17, November 1974.
89. A.M. Davis, E.A. Moustakas, "Analysis of Active RC Networks by Decomposition", IEEE Transactions on Circuits and Systems, Vol. CAS-27, May 1980.
90. M. Abraham, "Design of Butterworth-Type Transimpedance and Bootstrap-Transimpedance Preamplifiers for Fiber-Optic Receivers", IEEE Transactions on Circuits and Systems, Vol. CAS-29, June 1982.
91. M. Shigaki, S. Yokogawa, H. Kurihara, K. Yamada, "GaAs Monolithic DC-6.4GHz Variable-Gain Feedback Amplifier", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-35, October 1987.
92. J. Choma Jr., "Signal Flow Analysis of Feedback Networks", IEEE Transactions on Circuits and Systems, Vol. 37, April 1990.
93. S. Ben-Yaakov, "A Unified Approach to Teaching Feedback in Electronic Circuits Courses", IEEE Transactions on Education, Vol. 34, November 1991.
94. M. S. Gupta, "Power Gain in Feedback Amplifiers, a Classic Revisited", IEEE Transactions on Microwave Theory and Techniques, Vol. 40, May 1992.
95. P. J. Hurst, "A Comparison of Two Approaches to Feedback Circuit Analysis", IEEE Transactions on Education, Vol. 35, August 1992.
96. M. Vadipour, "Capacitive Feedback Technique for Wide-Band Amplifiers", IEEE Journal of Solid-State Circuits, Vol. 28, January 1993.
97. E. M. Cherry, "Input Impedance and Output Impedance of Feedback Amplifiers", IEE Proceedings on Circuits Devices and Systems, Vol. 143, August 1996.
98. G. Palumbo, J. Choma Jr., "An Overview of Analog Feedback Part I: Basic Theory", International Journal of Analog Integrated Circuits and Signal Processing, Kluwer, Vol. 17, November 1998.
99. G. Palumbo, J. Choma Jr., "An Overview of Analog Feedback Part II: Amplifier Configurations in Generic Device Technologies", International Journal of Analog Integrated Circuits and Signal Processing, Kluwer, Vol. 17, November 1998.
100. B. Nikolic, S. Marjanovic, "A General Method of Feedback Amplifier Analysis", ISCAS 98, IEEE International Symposium on Circuits and Systems, Monterey, CA, USA, May 1998.
101. P. Andreani, S. Mattison, "Extension of the Cochrun-Grabel Method to Allow for Mutual Inductances", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 46, April 1999.
102. G. Nowack, "Time-Delays in Microstructures and Stability-Problems in Ultra-High Speed Amplifiers for Optical WANs", Telsiks'99, IEEE International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services, Nis, Yugoslavia, October 1999.
103. W.-H. Ki, "Signal Flow Graph Analysis of Feedback Amplifiers", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 47, June 2000.
104. E. M. Cherry, "Feedback Amplifier Configurations", IEE Proceedings on Circuits Devices and Systems, Vol. 147, December 2000.
105. H. T. Russell, "A Loop-Breaking Method for the Analysis and Simulation of Feedback Amplifiers", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 49, August 2002.
106. F. Centurelli, R. Luzzi, M. Olivieri, A. Trifiletti, "A Bootstrap Technique for Wideband Amplifiers", IEEE Transactions on Circuits and Systems I, Vol. 49, October 2002.
107. J. Knorr, H. Zimmermann, "A Transmission-Line Approach for Modeling Feedback Resistors in Integrated Transimpedance Amplifiers", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 50, September 2003.
108. C. Seidl, J. Knorr, H. Zimmermann, "Simple Feedback Network for Bandwidth Enhancement of Transimpedance Amplifiers", IEE Electronic Letters, Vol. 39, December 2003.
109. M. R. Samadi, A. I. Karsilayan, J. Silva-Martinez, "Bandwidth Enhancement of Multi-Stage Amplifiers using Active Feedback", IEEE International Symposium on Circuits and Systems, Vancouver, Canada, May 2004.
110. J. L. Marrero, "Simplified Analysis of Feedback Amplifiers", IEEE Transactions on Education, Vol. 48, February 2005.
111. S. M. Potirakis, G. E. Alexakis, "An Accurate Calculation of Miller Effect on the Frequency Response and on the Input and Output Impedances of Feedback Amplifiers", IEEE Transactions on Circuits and

Systems – II, Vol. 52, August 2005.

112. C. Seidl, R. Swoboda, H. Zimmermann, “Optical Receiver IC with Capacitive Coupled Distributed Resistors Resulting in Effectively Reduced Parasitic Capacitance”, IEE Electronic Letters, Vol. 41, November 2005.
113. A. D’Amico, C. Falconi, G. Giustolisi, G. Palumbo, “Resistance of Feedback Amplifiers: A Novel Representation”, IEEE Transactions on Circuits and Systems – II: Express Briefs, Vol. 54, April 2007.
114. H.-Y. Huang, J.-C. Chien, L.-H. Lu, “A 10-Gb/s Inductorless CMOS Limiting Amplifier With Third-Order Interleaving Active Feedback”, IEEE Journal of Solid-State Circuits, Vol. 42, May 2007.

GBW enhancement circuit level techniques

115. J. R. Ragazzini, R. Randall, F. Russel, “Analysis of Problems in Dynamics by Electronic Circuits”, Proceedings of the IRE, Vol. 35, May 1947.
116. H. J. Carlin, D. C. Youla, “Network Synthesis with Negative Resistors”, Proceedings of the IRE, Vol. 49, May 1961.
117. E. M. Cherry, D. E. Hooper, “The Design of Wideband Transistor Feedback Amplifiers”, Proceedings of the IEE, Vol. 110, February 1963.
118. H. J. Carlin, “Singular Network Elements”, IEEE Transactions on Circuit Theory, Vol. CT-11, March 1964.
119. B. D. H. Tellegen, “On Nullators and Norators”, IEEE Transactions on Circuit Theory, Vol. CT-13, December 1966.
120. R. J. Widlar, “New Approaches for the Design of Monolithic Operational Amplifiers”, IEEE International Solid-State Circuits Conference, San Francisco, California, February 1969.
121. R. J. Widlar, “Design Techniques for Monolithic Operational Amplifiers”, IEEE Journal of Solid-State Circuits, Vol. SC-4, August 1969.
122. R. C. Jaeger, G. A. Hellwarth, “On the Performance of the Differential Cascode Amplifier”, IEEE Journal of Solid State Circuits, Vol. SC-8, April 1973.
123. A. A. Abidi, “On the Operation of Cascode Gain Stages”, IEEE Journal of Solid State Circuits, Vol. 23, December 1988.
124. G.-K. Chang, T. P. Liu, “Very High Speed Optical Receiver with a Cascadable Differential Action Feedback Driver”, US Patent 5345073, September 1994.
125. M. Sato, H. Shigematsu, Y. Inoue, T. Arai, K. Sawada, T. Takahashi, K. Makiyama, T. Hirose, “1.4-THz Gain-Bandwidth Product InP-HEMTs Preamplifier using an Improved Cherry-Hooper Topology”, IEEE GaAs Integrated Circuit Symposium, Monterey, California, October 2002.
126. C. D. Holdendried, M. W. Lynch, J. W. Haslett, “Modified CMOS Cherry-Hooper Amplifiers with Source Follower Feedback in 0.35 μ m Technology”, ESSCirc 2003, IEEE Proceedings of the European Solid-State Circuits Conference, Estoril, Portugal, September 2003.
127. C. D. Holdendried, J. W. Haslett, M. W. Lynch, “Analysis and Design of HBT Cherry-Hooper Amplifiers with Emitter-Follower Feedback for Optical Communications”, IEEE Journal of Solid State Circuits, Vol. 39, November 2004.
128. Y. Wu, “Wideband Constant-Gain Voltage Amplifier”, US Patent 6914484 B2, July 2005.
129. J. Abbott, C. Plett, J. W. M. Rogers, “A 15GHz, 1.8V, Variable-Gain, Modified Cherry-Hooper Amplifier”, CICC 2005, IEEE Custom Integrated Circuits Conference, San José, California, September 2005.
130. R. Krithivasan, J. Comeau, W.-M. L. Kuo, Y. Lu, J. D. Cressler, A. J. Joseph, “A 24GHz Broadband SiGe HBT Limiting Amplifier”, BCTM 2005, IEEE Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, Santa Barbara, California, December 2005.
131. C. Hermans, M. S. J. Steyaert, “A High-Speed 850-nm Optical Receiver Front-End in 0.18 μ m CMOS”, IEEE Journal of Solid State Circuits, Vol. 41, July 2006.

Stability of Time-Delay Systems

132. L. S. Pontryagin, “On the Zeros of Some Elementary Transcendental Functions”, American Mathematical Society Translation, Vol. 2, 1955.
133. C. S. Chang, “An Analytical Method for Obtaining the Root Locus with Positive and Negative Gain”,

- IEEE Transactions on Automatic Control, Vol. 10, January 1965.
134. I.-B. Huang, L. L.-C. Li, "Root Locus Determination of Linear Systems with Transport Lag", IEEE Transactions on Automatic Control, Vol. 12, October 1967.
135. I. H. Suh, Z. Bien, "A Root-Locus Technique for Linear Systems with Delay", IEEE Transactions on Automatic Control, Vol. AC-27, February 1982.
136. J. S. Freudenberg, D. P. Louze, "Right Half Plane Poles and Zeros and Design Tradeoffs in Feedback Systems", IEEE Transactions on Automatic Control, Vol. AC-30, June 1985.
137. J. S. Freudenberg, D. P. Louze, "A Sensitivity Tradeoff for Plants with Time Delay", IEEE Transactions on Automatic Control, Vol. AC-32, February 1987.
138. H. Chapellat, M. Mansour, S. P. Bhattacharyya, "Elementary Proofs of Some Classical Stability Criteria", IEEE Transactions on Education, Vol. 3, August 1990.
139. V. L. Kharitonov, A. P. Zhabko, "Robust Stability of Time-Delay Systems", IEEE Transactions on Automatic Control, Vol. 39, December 1994.
140. W. K. Ho, C. C. Hang, J. H. Zhou, "Performance and Gain and Phase Margins of Well-Known PI Tuning Formulas", IEEE Transactions on Control Systems Technology, Vol. 3, June 1995.
141. W. K. Ho, C. C. Hang, L. S. Cao, "Tuning of PID Controllers Based on Gain and Phase Margin Specifications", Elsevier, Automatica, Vol. 31, 1995.
142. L. H. Keel, S. P. Bhattacharyya, "Phase Properties of Hurwitz polynomials", IEEE Transactions on Automatic Control, Vol. 41, May 1996.
143. W. K. Ho, O. P. Gan, E. B. Tay, E. L. Ang, "Performance and Gain and Phase Margins of Well-Known PID Tuning Formulas", IEEE Transactions on Control Systems Technology, Vol. 4, July 1996.
144. M.-T. Ho, A. Datta, S. P. Bhattacharyya, "A New Approach to Feedback Stabilization", IEEE Proceedings of the 35th Conference on Decision and Control, Kobe, Japan, December 1996.
145. W. K. Ho, C. C. Hang, J. Zhou, "Self-Tuning PID Control of a Plant with Under-Damped Response with Specifications on Gain and Phase Margins", IEEE Transactions on Control Systems Technology, Vol. 5, July 1997.
146. M.-T. Ho, A. Datta, S. P. Bhattacharyya, "An Elementary Derivation of the Routh-Hurwitz Criterion", IEEE Transactions on Automatic Control, Vol. 43, March 1998.
147. M.-T. Ho, A. Datta, S. P. Bhattacharyya, "An Extension of the Generalized Hermite-Biehler Theorem: Relaxation of Earlier Assumptions", IEEE Proceedings of American Control Conference, Philadelphia, Pennsylvania, June 1998.
148. F. M. Asl, A. G. Ulsoy, "Analytical Solution of a System of Homogeneous Delay Differential Equations via the Lambert Function", IEEE Proceedings of the American Control Conference, Chicago, Illinois, June 2000.
149. G. J. Silva, A. Datta, S. P. Bhattacharyya, "Stabilization of Time Delay Systems", IEEE Proceedings of the American Control Conference, Chicago, Illinois, June 2000.
150. G. J. Silva, A. Datta, S. P. Bhattacharyya, "Stabilization of First-order Systems with Time Delay using the PID controller", IEEE Proceedings of the American Control Conference, Arlington, VA, June 2001.
151. G. J. Silva, A. Datta, S. P. Bhattacharyya, "Determination of Stabilizing Feedback Gains for Second-order Systems with Time Delay", IEEE Proceedings of the American Control Conference, Arlington, VA, June 2001.
152. G. J. Silva, A. Datta, S. P. Bhattacharyya, "PI stabilization of first-order systems with time delay", Elsevier, Automatica, Vol. 37, December 2001.
153. Y. Chen, K. L. Moore, "Analytical stability bound for delayed second-order systems with repeating poles using Lambert function W", Elsevier, Automatica, Vol. 38, May 2002.
154. J. Silva, A. Datta, S. P. Bhattacharyya, "New Results on the Synthesis of PID Controllers", IEEE Transactions on Automatic Control, Vol. 47, February 2002.
155. H. Xu, A. Datta, S. P. Bhattacharyya, "PID Stabilization of LTI Plants with Time-Delay", Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Hawaii, December 2003.
156. Y.-C. Cheng, C. Hwang, "Use of the Lambert W function for time-domain analysis of feedback fractional delay systems", IEEE Proceedings on Control Theory and Applications, Vol. 153, March 2006.

Transistor modeling

157. P. R. Drouilhet, "Predictions Based on the Maximum Oscillator Frequency of a Transistor", IRE Transactions on Circuit Theory, Vol. CT-2, June 1955.
158. R. L. Pritchard, "Electric-Network Representation of Transistors – A Survey", IRE Transactions on Circuit Theory, Vol. CT-2, June 1956.
159. R. P. Abraham, "Transistor Behavior at High Frequencies", IRE Transactions on Electronic Devices, Vol. ED-7, January 1960.
160. J. J. Kelly, M. S. Ghausi, "Poles of the Alpha Drift Transistors", IEEE Transactions on Circuit Theory, Vol. 12, December 1965.
161. J. E. Solomon, G. R. Wilson, "A Highly Desensitized, Wide-Band Monolithic Amplifier", IEEE Journal of Solid-State Circuits, Vol. SC-1, September 1966.
162. M. B. Steer, R. J. Trew, "High-Frequency Limits of Millimeter-Wave Transistors", IEEE Electron Device Letters, Vol. EDL-7, November 1986.
163. W. R. Curtice, "GaAs MESFET Modeling and Nonlinear CAD", IEEE Transactions on Microwave Theory and Techniques, Vol. 36, February 1988.
164. W. R. Curtice, "Intrinsic GaAs MESFET Equivalent Circuit Models Generated From Two-Dimensional Simulations", IEEE Transactions on Computer Aided Design, Vol. 8, April 1989.
165. D. A. Teeter, J. R. East, R. K. Mains, G. I. Haddad, "Large-Signal Numerical and Analytical HBT Models", IEEE Transactions on Electron Devices, Vol. 40, May 1993.
166. K. Nummila, A. A. Ketterson, I. Adesida, "Delay Time Analysis for Short Gate-Length GaAs MESFETs", Elsevier Solid-State Electronics, Vol. 38, 1995.
167. A. Samelis, D. Pavlidis, "DC to High-Frequency HBT-Model Parameter Evaluation Using Impedance Block Conditioned Optimization", IEEE Transactions on Microwave Theory and Techniques, Vol. 45, June 1997.
168. G. I. Haddad, R. J. Trew, "Microwave Solid-State Active Devices", IEEE Transactions on Microwave Theory and Techniques, Vol. 50, March 2002.
169. I. Kwon, M. Je, K. Lee, H. Shin, "A Simple and Analytical Parameter-Extraction Method of a Microwave MOSFET", IEEE Transactions on Microwave Theory and Techniques, Vol. 50, June 2002.
170. M. W. Dvorak, C. R. Bolognesi, "On the Accuracy of Direct Extraction of the Heterojunction-Bipolar-Transistor Equivalent-Circuit Model Parameters C_{π} , CBC, and RE", IEEE Transactions on Microwave Theory and Techniques, Vol. 51, June 2003.
171. S. V. Cherepko, J. C. M. Hwang, "Implementation of Nonquasi-Static Effects in Compact Bipolar Transistor Models", IEEE Transactions on Microwave Theory and Techniques, Vol. 51, December 2003.
172. R. J. Trew, "High-Frequency Solid-State Electronic Devices", IEEE Transactions on Electron Devices, Vol. 52, May 2005.
173. D. Dousset, A. Issaoun, F. M. Ghannouchi, A. B. Kouki, "Wideband Closed-Form Expressions for Direct Extraction of HBT Small-Signal Parameters for all Amplifier Bias Classes", IEE Proceedings on Circuits, Devices and Systems, Vol. 152, October 2005.

Transmission Lines Theory and Practice

174. H. A. Wheeler, "Formulas for the Skin Effect", Proceedings of the IRE, Vol. 30, September 1942.
175. S. B. Cohn, "Characteristic Impedance of the Shielded-Strip Transmission Line", IRE Transactions on Microwave Theory and Techniques, Vol. 2, July 1954.
176. S. B. Cohn, "Problems in Strip Transmission Lines", IRE Transactions on Microwave Theory and Techniques, Vol. 3, March 1955.
177. S. B. Cohn, "Shielded Coupled-Strip Transmission Line", IRE Transactions on Microwave Theory and Techniques, Vol. 3, October 1955.
178. S. B. Cohn, "Characteristic Impedances of Broadside-Coupled Strip Transmission Lines", IRE Transactions on Microwave Theory and Techniques, Vol. 8, November 1960.
179. H. A. Wheeler, "Transmission-Line Properties of a Parallel Wide Strips by a Conformal-Mapping Approximation", IEEE Transactions on Microwave Theory and Techniques, Vol. 12, May 1964.
180. H. A. Wheeler, "Transmission-Line Properties of Parallel Strips Separated by a Dielectric Sheet", IEEE

- Transactions on Microwave Theory and Techniques, Vol. 13, March 1965.
181. H. E. Stinehelfer, "An Accurate Calculation of Uniform Microstrip Transmission Lines", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-16, July 1968.
182. P. J. Tu, "A Computer-Aided Design of a Microwave Delay Equalizer", IEEE Transactions on Theory and Techniques, Vol. MTT-17, August 1969.
183. R. Sato, "A Design Method for Meander-Line Networks Using Equivalent Circuit Transformations", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-19, May 1971.
184. H. A. Wheeler, "Transmission-Line Properties of a Strip on a Dielectric Sheet on a Plane", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-25, August 1977.
185. H. A. Wheeler, "Transmission-Line Properties of a Strip Line Between Parallel Planes", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-26, November 1978.
186. W. H. Haydl, "Properties of Meander Coplanar Transmission Lines", IEEE Microwave and Guided Wave Letters, Vol. 2, November 1992.
187. G. B. Stracca, "A Simple Evaluation of Losses in Thin Microstrips", IEEE Transactions on Microwave Theory and Techniques, Vol. 45, February 1997.
188. B. J. Rubin, B. Singh, "Study of Meander Line Delay in Circuit Boards", IEEE Transactions on Microwave Theory and Techniques, Vol. 48, September 2000.
189. F. Schnieder, W. Heinrich, "Model of Thin-Film Microstrip Line for Circuit Design", IEEE Transactions on Microwave Theory and Techniques, Vol. 49, January 2001.
190. H. Lee, J. Kim, "Unit Cell Approach to Full-Wave Analysis of Meander Delay Line Using FDTD Periodic Structure Modeling Method", IEEE Trans. on Advanced Packaging, Vol. 25, May 2002.
191. T. Sudo, J. Kudo, Y. Ko, K. Ito, "Experimental Characterization and Numerical Modeling Approach of Meander Delay Lines", EMC 2002, IEEE International Symposium on Electromagnetic Compatibility, Minneapolis, Minnesota, August 2002.
192. C. H. Doan, S. Emami, D. A. Sobel, A. M. Niknejad, R. W. Brodersen, "Design Considerations for 60 GHz CMOS Radios", IEEE Communications Magazine, December 2004.
193. C. H. Doan, S. Emami, A. M. Niknejad, R. W. Brodersen, "Millimeter-Wave CMOS Design", IEEE Journal of Solid-State Circuits, Vol. 40, January 2005.
194. B. Razavi, "A 60 GHz CMOS Receiver Front-End", IEEE Journal of Solid-State Circuits, Vol. 41, January 2006.

Active Delay Synthesis

195. D. O. Pederson, G. H. Wilson, "Dominant Zero and Excess Phase of a Hurwitz Polynomial", IRE Transactions on Circuit Theory, Vol. 11, March 1964.
196. R. Gregorian, G. C. Temes, "Design Techniques for Digital and Analog All-Pass Circuits", IEEE Transactions on Circuits and Systems, Vol. CAS-25, December 1978.
197. J. Choma Jr., "A Generalized Bandwidth Estimation Theory for Feedback Amplifiers", Transactions on Circuits and Systems, Vol. CAS-31, October 1984.
198. K. Bult, H. Wallinga, "A CMOS Analog Continuous-Time Delay Line with Adaptive Delay-Time Control", IEEE Journal of Solid-State Circuits, Vol. 23, June 1988.
199. M. G. Johnson, E. L. Hudson, "A Variable Delay Line PLL for CPU-Coprocessor Synchronization", IEEE Journal of Solid-State Circuits, Vol. 23, October 1988.
200. G. Kim, M.-K. Kim, B.-S. Chang, W. Kim, "A Low-Voltage, Low-Power CMOS Delay Element", IEEE Journal of Solid-State Circuits, Vol. 31, July 1996.
201. A. Borjak, P. P. Monteiro, J. J. O'Reilly, I. Darwazeh, "High-Speed Generalized Distributed-Amplifier Based Transversal-Filter Topology for Optical Communication Systems", IEEE Transactions on Microwave Theory and Techniques, Vol. 45, August 1997.
202. I. M. Filanovsky, "One Method of Synthesis of Time Delay Circuits", MWSCAS 99, IEEE Midwest Symposium on Circuits and Systems, Las Cruces, New Mexico, August 1999.
203. B. Analui, A. Hajimiri, "Statistical Analysis of Integrated Passive Delay Lines", CICC 2003, IEEE Custom Integrated Circuits Conference, San José, California, September 2003.
204. M. Maymandi-Nejad, M. Sachdev, "A Digitally Programmable Delay Element: Design and Analysis", IEEE Transactions on Very Large Scale Integration Systems, Vol. 11, October 2003.
205. H. Wu, J. A. Tierno, S. Gowda, J. A. Kash, A. Hajimiri, "Integrated Transversal Equalizers in High-

- Speed Fiber-Optic Systems”, IEEE Journal of Solid-State Circuits, Vol. 38, December 2003.
206. H. Yoo, D. Graham, D. V. Anderson, P. Hasler, “ C^4 Band-Pass Delay Filter for Continuous-Time Subband Adaptive Tapped-Delay Filter”, ISCAS 2004, IEEE International Symposium on Circuits and Systems, Vancouver, Canada, May 2004.
 207. J. Buckwalter, A. Hajimiri, “An Active Delay and the Delay Reference Loop”, RFIC 2004, IEEE Radio Frequency Integrated Circuits Symposium, Fox Worth, Texas, June 2004.
 208. C. Pelard, E. Gebara, A. J. Kim, M. G. Vrazel, F. Bien, Y. Hur, M. Maeng, S. Chandramouli, C. Chun, S. Bajekal, S. E. Ralph, B. Schmukler, V. M. Hietala, J. Laskar, “Realization of Multigigabit Channel Equalization and Crosstalk Cancellation Integrated Circuits”, IEEE Journal of Solid-State Circuits, Vol. 39, October 2004.
 209. P. Ziska, J. Vrbata, “Method for Design of Analog Group Delay Equalizers”, ISCAS 2006, IEEE International Symposium on Circuits and Systems, Kos, Greece, May 2006.
 210. S. M. Kashmiri, S. A. P. Haddad, W. A. Serdijn, “High-Performance Analog Delays: Surpassing Bessel-Thomson by Padé-Approximated Gaussians”, ISCAS 2006, IEEE International Symposium on Circuits and Systems, Kos, Greece, May 2006.
 211. R. Tiruvuru, S. Pavan, “Transmission Line based FIR Structures for High Speed Adaptive Equalization”, ISCAS 2006, IEEE Inter. Symp. on Circuits and Systems, Kos, Greece, May 2006.
 212. T. Lovitt, C. Plett, J. Rogers, “A 0.13 μ m CMOS Delay Cell for 40 Gb/s FFE Equalization”, ISCAS 2006, IEEE International Symposium on Circuits and Systems, Kos, Greece, May 2006.

Current-Mode & Current Conveyors

213. G. E. Moore, “Cramming More Components onto Integrated Circuits”, Proceedings of the IEEE, Vol. 86, January 1998. – Reprinted from, G. E. Moore, “Cramming More Components onto Integrated Circuits”, Electronics, April 1965.
214. A. F. Arbel, “Multistage Transistorized Current Modules”, IEEE Transactions on Circuit Theory, Vol. CT-13, September 1966.
215. K. C. Smith, A. Sedra, “The Current Conveyor – A new Circuit Building Block”, Proceedings of the IEEE, Vol. 56, August 1968.
216. A. Sedra, K. C. Smith, “A Second-Generation Current Conveyor and its Applications”, IEEE Transactions on Circuit Theory, Vol. CT-17, February 1970.
217. R. H. Dennard, F. H. Gaensslen, L. Kuhn, H. N. Yu, “Design of Micron MOS Switching Devices”, IEDM, International Electron Devices Meeting, Washington D.C., U.S.A., December 1972.
218. R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, A. LeBlanc, “Ion Implanted MOSFETs with Very Short Channel Lengths”, IEDM, International Electron Devices Meeting, Washington D.C., U.S.A., December 1973.
219. R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, A. LeBlanc, “Design of Ion-Implanted MOSFET’s with Very Small Physical Dimensions” IEEE Journal of Solid-State Curcuits, Vol. SC-9, October 1974.
220. G. E. Moore, “Progress in Digital Integrated Electronics”, IEDM, International Electron Devices Meeting, Washington D.C., U.S.A., December 1975.
221. G. Baccarani, M. R. Wordeman, R. H. Dennard, “Generalized Scaling Theory and its Application to a $\frac{1}{4}$ Micrometer MOSFET Design”, IEEE Transactions on Electron Devices, Vol. ED-31, April 1984.
222. B. Wilson, “Performance Analysis of Current Conveyors”, IEE Electronics Letters, Vol. 25, November 1989.
223. A. S. Sedra, G. W. Roberts, F. Gohh, “The Current Conveyor: History, Progress and new Results”, IEE Proceedings, Vol. 137, April 1990.
224. T. Laopoulos, S. Siskos, M. Bafleur, P. Givelin, “CMOS Current Conveyor”, IEE Electronics Letters, Vol. 28, November 1992.
225. M. C. H. Cheng, C. Toumazou, “3V MOS Current Conveyor Cell for VLSI Technology”, IEE Electronics Letters, Vol. 29, February 1993.
226. A. Fabre, H. Barthelemy, “Composite Second-Generation Current Conveyor with Reduced Parasitic Resistance”, IEE Electronics Letters, Vol. 30, March 1994.
227. A. Fabre, “Third-Generation Current Conveyor: A New Helpful Active Element”, IEE Electronics

- Letters, Vol. 31, March 1995.
228. O. Oliaei, P. Loumeau, "Current-Mode Class AB Design Using Floating Voltage-Source", IEE Electronics Letters, Vol. 32, August 1996.
229. E. Bruun, "Analysis of the Noise Characteristics of CMOS Current Conveyors", Journal of Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers, Vol. 12, January 1997.
230. A. Arbel, "Towards a Perfect CMOS CCII", International Journal of Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers, Vol. 12 February 1997.
231. O. Oliaei, J. Porte, "Compound Current Conveyor (CCII⁺ and CCII⁻)", IEE Electronics Letters, Vol. 33, February 1997.
232. A. Fabre M. Alami, "A Precise Macromodel for Second Generation Current Conveyors", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 44, July 1997.
233. G. Palumbo, S. Pennisi, "A Technique for the Reduction of the Input Resistance of Current-Mode Circuits", ICECS 1998, IEEE International Conference on Electronic Circuits and Systems, Lisboa Portugal, September 1998.
234. T. Ruotsalainen, "A Current-Mode Gain-Control Scheme with Constant Bandwidth and Propagation Delay for a Transimpedance Preamplifier", IEEE Journal of Solid-State Circuits, Vol. 34, February 1999.
235. G. Palmisano, G. Palumbo, S. Pennisi, "Design Strategies for Class A CMOS CCII's", International Journal of Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers, Vol. 19 April 1999.
236. I. A. Awad, A. M. Soliman, "New CMOS Realization of the CCII-", IEEE Transactions on Circuits and Systems – II: Analog and Digital Signal Processing, Vol. 46, April 1999.
237. E. Bruun, "On Dynamic Range Limitations of CMOS Current Conveyors", ISCAS 1999, IEEE International Symposium on Circuits and Systems, Orlando, Florida USA May 1999.
238. K. Phang, D. A. Johns, "A CMOS Optical Preamplifier for Wireless Infrared Communications", IEEE Transactions on Circuits and Systems - II, Vol. 46, July 1999.
239. H. Schmid, "Approximating the Universal Active Element", IEEE Transactions on Circuits and Systems – II: Analog and Digital Signal Processing, Vol. 47, November 2000.
240. F. Seguin, A. Fabre, "2GHz Controlled Current Conveyor in Standard 0.8 μ m BiCMOS Technology", IEE Electronics Letters, Vol. 37, March 2001.
241. F. Seguin, A. Fabre, "New Second Generation Current Conveyor with Reduced Parasitic Resistance and Bandpass Filter Application", IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, Vol. 48, June 2001.
242. G. E. Moore, "No Exponential is Forever: But "Forever" Can Be Delayed!", ISSCC, International Solid-State Circuits Conference, San Francisco, U.S.A., February 2003.
243. A. S. Hou, "A Wide Bandwidth Isolation Amplifier Design Using Current Conveyors", Journal of Analog Integrated Circuits and Signal Processing, Vol. 40, July 2004.
244. F. Seguin, B. Godara, F. Alicalapa, A. Fabre, "2.2GHz All-n-p-n Second-Generation Controlled Conveyor in Pseudoclass AB Using 0.8- μ m BiCMOS Technology", IEEE Transactions on Circuits and Systems – II: Express Briefs, Vol. 51, July 2004.
245. B. Godara, A. Fabre, "Versatile Wideband Impedance Matching Circuit Based on Current Conveyors", IEE Electronics Letters, Vol. 43 March 2007.

General references

246. R. M. Fano, "Theoretical Limitations on the Broad Band Matching of Arbitrary Impedances", PhD Dissertation, Massachusetts Institute of Technology, 1947.
247. F. E. Terman, "Electronic and Radio Engineering", 4thED., McGraw-Hill Kogakusha, 1955.
248. H. W. Bode, "Network Analysis and Feedback Amplifier Design", D. Van Nostrand Company Inc., 1956.
249. L. S. Pontryagin, "Ordinary Differential Equations", Addison-Wesley Publishing Company, Inc., 1962.
250. M. Abramowitz, I. A. Stegun, "Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical tables", Dover publications, 1965.
251. F. F. Kuo, "Network Analysis and Synthesis", 2ndED., John Wiley & Sons, 1966.
252. J. Millman, C. C. Halkias, "Integrated Electronics", McGraw-Hill, 1971.

253. S.M. Sze, "Physics of Semiconductor Devices", 2^oED., John Wiley & Sons, 1981.
254. M. L. Boas, "Mathematical Methods in the Physical Sciences", 2^oED., John Wiley & Sons, 1983.
255. J. D. Kraus, "Electromagnetics", 3^oED., McGraw-Hill International Editions, 1984.
256. J. Spanier, K. B. Oldham, "An Atlas of Functions", Hemisphere Publishing Corporation, 1987.
257. J. J. D'Azzo, C.H. Houpis, "Linear Control System Analysis and Design Conventional and Modern", McGraw-Hill International Editions, 3^oED, 1988.
258. C. W. Davidson, "Transmission Lines for Communications with CAD programs", MacMillan 1989.
259. C. Toumazou, F. J. Lidgey, D. G. Haigh, "Analogue IC Design: The Current-Mode Approach", Peter Peregrinus, 1990.
260. S.Y. Liao, "Microwave Devices and Circuits", 3^oED., Prentice-Hall, 1990.
261. B. C. Wadell, "Transmission Line Design Handbook", Artech House 1991.
262. F. Ali, A. Gupta, "HEMTs & HBTs, Devices, Fabrication, and Circuits", Artech House 1991.
263. C. Toumazou, N. C. Battersby, J. B. Hughes, "Switched-Currents: An Analogue Technique for Digital Technology", IEE Publishing, 1993.
264. P. R. Gray, R. G. Meyer, "Analysis and Design of Analog Integrated Circuits", 3^oED., John Wiley & Sons 1993.
265. R. J. Wiegink, E. Seevinck, "Analysis and Synthesis of MOS Translinear Circuits", Kluwer Academic Publishers, 1993.
266. T. T. Y. Wong, "Fundamentals of Distributed Amplification", Artech House, 1993.
267. W. K. Chen, "The Circuits and Filters Handbook", CRC Press, 1995.
268. G. A. Baker Jr., P. G.-Morris, "Padé Approximations – 2^o edition", Cambridge University Press, 1996.
269. J. C. Freeman, "Fundamentals of Microwave Transmission Lines", John Wiley & Sons 1996.
270. J. Graeme, "Photodiode Amplifiers: OP AMP Solutions", McGraw-Hill, 1996.
271. S. B. Alexander, "Optical Communication Receiver Design", SPIE, 1997.
272. K. Kano, "Semiconductor Devices", Prentice Hall, 1998.
273. G. Palmisano, G. Palumbo, S. Pennisi, "CMOS Current Amplifiers", Kluwer Academic Publishers, 1999.
274. W. Liu, "Fundamentals of III-V Devices, HBTs, MESFETs, and HFETs/HEMTs", John Wiley & Sons, 1999.
275. J. S. Yuan, "SiGe, GaAs and InP Heterojunction Bipolar Transistors", John Wiley & Sons, 1999.
276. C. Toumazou, G. Moschytz, B. Gilbert, "Trade-Offs in Analog Circuit Design – The Designer Companion", Kluwer Academic Publishers, 2002.
277. G. Palumbo, S. Pennisi, "Feedback Amplifiers Theory and Design", Kluwer Academic Publishers, 2002.
278. K. Koli, K. Halonen, "CMOS Current Amplifiers, Speed versus Nonlinearity", Kluwer Academic Publishers, 2002.
279. B. Razavi, "Design of Integrated Circuits for Optical Communications", McGraw-Hill, 2003.
280. C. J. M. Verhoeven, A. Staveren, G. L. E. Monna, M. H. L. Kouwenhoven, E. Yildiz, "Structured Electronic Design – Negative-Feedback Amplifiers", Kluwer Academic Publishers, 2003.
281. G. Ferri, N. C. Guerrini, "Low-Voltage Low-Power CMOS Current Conveyors", Kluwer Academic Publishers, 2003.
282. K. Gu, V. L. Kharitonov, J. Chen, "Stability of Time-Delay Systems", Birkhäuser, 2003.
283. D. Hinrichsen, A. J. Pritchard, "Mathematical Systems Theory I: Modelling, State Space Analysis, Stability and Robustness", Springer-Verlag, 2005.
284. E. Säckinger, "Broadband Circuits for Optical Fiber Communications", John Wiley & Sons, 2005.
285. G. J. Silva, A. Datta, S. P. Bhattacharya, "PID Controllers for Time-Delay Systems", Birkhäuser, 2005.
286. P. Staric, E. Margan, "Wideband Amplifiers", Springer 2006.

Original contributions

287. L. N. Alves, R. L. Aguiar, "A Differential Current-Conveyor Based Buffer for High-Bandwidth, Low-Signal Applications", ICECS 1999, IEEE International Conference of Electronic Circuits and Systems, Paphos, Cyprus, September 1999.
288. L. N. Alves, R. L. Aguiar, "Maximizing Bandwidth in CCII for Wireless Optical Applications", ICECS

- 2001, IEEE International Conference on Electronic Circuits and Systems, St. Julians, Malta, September 2001.
289. L. N. Alves, R. L. Aguiar, D. M. Santos, "Bandwidth Aspects in Second Generation Current Conveyors", DCIS 2001 – XVI Conference on Design of Circuits and Integrated Systems. Porto, Portugal, November 2001.
290. L. N. Alves, R. L. Aguiar, "Noise Performance of Classical Current Mirrors", ICECS 2002, IEEE International Conference on Electronic Circuits and Systems, Dubrovnik, Croatia, September 2002.
291. L. N. Alves, R. L. Aguiar, "Frequency Behavior of Classical Current Mirrors", ICECS 2002, IEEE International Conference on Electronic Circuits and Systems, Dubrovnik, Croatia, September 2002.
292. L. N. Alves, R. L. Aguiar, D. M. Santos, "Bandwidth Aspects in Second Generation Current Conveyors", *Journal of Analog Integrated Circuits and Signal Processing*, Kluwer Academic Publishers, Vol. 33, November, 2002.
293. L. N. Alves, R. L. Aguiar, "Design Techniques for High Performance Optical Wireless Front-Ends", *Proceedings of the Conference on Telecommunications - ConfTele 2003*, Aveiro, Portugal, April, 2003.
294. L. N. Alves, R. L. Aguiar, "Extending the Laplace Expansion Method to the Frequency Response Analysis of Active RLC circuits", ICECS 2003, IEEE International Conference on Electronic Circuits and Systems, Sharjah, United Arab Emirates, December 2003.
295. L. N. Alves, R. L. Aguiar, "A New Method to Improve the Impedance of CC-II's X Input", ICECS 2004, IEEE International Conference on Electronic Circuits and Systems, Tel-Aviv, Israel, December 2004.
296. L. N. Alves, R. L. Aguiar, "On the Effect of Time Delays in Negative Feedback Amplifiers", ISCAS 2005, IEEE International Symposium on Circuits and Systems, Kobe, Japan, May 2005.
297. L. N. Alves, R. L. Aguiar, "On the Usage of Delayed-Feedback in Amplifiers", ECCTD 2005, European Conference on Circuit Theory and Design 2005, Cork, Ireland, September 2005
298. N. Martins, L. Barbosa, L. N. Alves, R. L. Aguiar, "Bandwidth Boosting in Transimpedance Amplifiers for Optical LANs, using Delayed Feedback", ICECS 2005, IEEE International Conference on Electronic Circuits and Systems, Ganmarth, Tunisia, December 2005.
299. L. N. Alves, L. Barbosa, R. L. Aguiar, "General Model for the Deployment of Time-delay Elements in Transistorized Electronic Circuits", IEEE International Conference on Electronic Circuits and Systems, Nice, France, December 2005.
300. L. N. Alves, L. Barbosa, E. A. L. Macedo, R. L. Aguiar, "General Model for Delayed Feedback and its Application to Transimpedance Amplifier's Bandwidth Optimization", ISCAS 2006, IEEE International Symposium on Circuits and Systems, Kos, Greece, May 2006.
301. L. N. Alves, R. L. Aguiar, "A Time-Delay Technique to Improve GBW on Negative Feedback Amplifiers", *Journal of Circuit Theory and Applications*, Wiley, June 2008.